THE UNIVERSITY *of* EDINBURGH

DISRUPTING THE INCENP-AURORA B INTERACTION WITH GENETICALLY-ENCODED CYCLIC PEPTIDES

FLORENCE GOHARD

THESIS PRESENTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

WELLCOME TRUST CENTRE FOR CELL BIOLOGY
INSTITUTE OF CELL BIOLOGY
UNIVERSITY OF EDINBURGH
APRIL 2014

DECLARATION

I hereby declare that this thesis is composed by myself and that the work presented therein is my own, unless stated otherwise.

Florence Gohard

April 2014

ACKNOWLEDGEMENTS

Firstly, I wish to extend my thanks to my supervisor Bill Earnshaw and my second supervisors, Mike Tyers and Malcolm Walkinshaw, for their guidance, support and advice throughout the course of this project and beyond. I thank Daniel St-Cyr and Samir Bouayad-Gervais for kindly synthesizing hit peptides for this project. I am grateful to the former mentors and supervisors who have encouraged me to pursue this path, as I am to friends and colleagues from the Earnshaw lab and beyond for their advice and helpful discussions.

I thank the Wellcome Trust for funding this research and my PhD studies. Moreover, I wish to jointly thank the Trust, the Academy of Medical Sciences and their respective policy teams for their Policy Internship scheme, which I had the pleasure of participating in during my PhD.

Warm thanks to my friends, old and new – your friendship and good humour has seen me through the low points of this adventure and made the high ones all the more enjoyable. I am grateful to Dave Seaton for humouring me and agreeing to be my thesis beard surrogate. For their long-standing friendship, I wish to personally thank Al Needham, John Ferguson, Caroline Pugh, and most importantly Phillip Quirie - you are my best friend and you have kept me grounded throughout this. I also wish to thank Stephen Strowse and Fernanda Loureiro Goulart for their support during my write up.

I extend my heartfelt thanks to my family for the love, support, encouragement and patience they have always given me. Thank you for never holding me back, and always welcoming me home with open arms.

DEDICATION

For her reflections on the wealth of opportunities we now enjoy and her encouragement to fearlessly pursue these, this thesis is dedicated to the memory of my maternal grandmother, Thérèse Dru.

Pour ses réflexions sur la richesse des opportunités dont nous jouissons aujourd'hui et pour m'avoir encouragée à poursuivre celles-ci sans crainte, je dédie cette thèse à la mémoire de ma grand-mère maternelle, Thérèse Dru.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACA	Anti-centromere antibody
APC/C	anaphase promoting complex/cyclosome
ATP	adenosine triphosphate
BIR	baculovirus inhibitor of apoptosis
BSA	bovine serum albumin
C-terminal/us	carboxy-terminal/-terminus
CENP	centromere protein
CMV	Cytomegalo virus
CPC	centromeric passenger complex
CPC	chromosomal passenger complex
CPP	cell penetrating peptide
DAPI	4',6'-diamidino-2-phenylindole
ddH_20	double distilled water
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
Drosophila	*Drosophila melanogaster*
E coli	*Escherichia coli*
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
FBS	fetal bovine serum
GFP	green fluorescent protein
IAP	inhibitor of apoptosis
IL	interleukin
INCENP	inner centromere protein
IRES	internal ribosome entry site
kb	Kilobase
kDa	Kilodalton
KMN	KNL-1-Mis12-Ndc80
LB	Luria Bertani
mg	milligram
ml	milliliter
mM	Molar
MCC	mitotic checkpoint complex
N-terminal/us	amino-terminal/-terminus
NEB	New England Biolabs
PBS	Phosphate buffered saline
PBS-t	PBS with added Tween-20
PCR	Polymerase chain reaction

PPI	protein-protein interaction
RCF	relative centrifugal force
RNA	ribonucleic acid
RPM	revolutions per minute
SDS	sodium docecyl sulfate
SDS-PAGE	SDS polyacrylamide electrophoresis
SICLOPPS	split-intein ligation of proteins and peptides
Ssp	*Synechocystis* species PCC6803
TAE	Tris acetate EDTA
TD-60	telophase disk 60
TSS	threonine-serine-serine
UV	ultraviolet
v/v	volume/volume
w/v	weight/volume
WT	wild type
Xenopus	*Xenopus laevis*
µg	microgram
µl	microlitre
µM	micromolar

TABLE OF CONTENTS

ABSTRACT

The chromosome passenger complex (CPC) is an essential mitotic regulator with key roles in mitotic processes such as chromosome condensation, spindle dynamics, chromosome bi-orientation, the spindle checkpoint and cytokinesis. The Aurora B kinase is the CPC's catalytic subunit. Its targeting and activation are dependent on interactions with the other components of the complex: inner centromere protein (INCENP), survivin and borealin/Dasra B. INCENP serves both as a scaffolding subunit for the CPC as a whole and as an activator of Aurora B via its highly conserved INbox domain.

Aurora B is a putative anti-cancer target; several inhibitors of the kinase are currently in clinical trials. All these are ATP-analogues targeting the kinase active site. The protein-protein interaction between Aurora B and the INCENP INbox is also essential for CPC function. Earlier studies have demonstrated that INCENP INbox mutants unable to bind and/or activate Aurora B cannot rescue lethality in the absence of endogenous INCENP.

The first goal of this study was to test the *in vivo* effects of disrupting the interaction between endogenous wild type INCENP and Aurora B. For this, a cell-based CPC function assay was developed in HeLa cells. Using this assay, I show that expression of soluble INbox in HeLa cells produces a significant increase in multinucleated and micronucleated cells: both effects consistent with Aurora B loss of function. Expression of soluble INbox bearing the mutations W845G and/or F881A does not elicit this effect suggesting that those mutants cannot bind to Aurora B and occlude INCENP binding. The result concerning the F881A mutant contrasts with earlier reports that equivalent mutants could bind, but not activate, Aurora B. Expression of an

INbox mutant lacking the C-terminal TSS motif reported to be involved in Aurora B activation but not binding has effects similar to those of the wild type INbox.

Using the INbox/Aurora B interaction as a model, a secondary goal of this study was to develop and evaluate a novel approach to identify small peptides capable of dissociating intracellular protein-protein interactions. For this, a library of small (5-9 residues long) circular peptides (CPs) mimicking the INbox was generated using the split intein circular ligation of proteins and peptides (SICLOPPS) methodology and assayed using the cell-based CPC function assay. Over two successive rounds of screening, a small number of CPs were identified that caused a significant increase in rates of multinucleated and micronucleated cells. Although statistically significant, these increases were very modest. Furthermore, due to high heterogeneity in SICLOPPS processing efficiencies, it was not practicable to compare the effects of different peptides side-by-side by transfection. The level of variation in processing efficiency – thus, CP production – was unexpectedly high and puts into question the functional complexity of more commonly used combinatorial cyclic peptide libraries derived using current SICLOPPS methodology.

The results of this study are divided into three sections. The first is a methods section concerning the testing of SICLOPPS in HeLa cells and the development of a cell-based CPC function assay. In the second, the effects of expressing soluble INbox and mutants thereof in HeLa cells are presented. The final results section presents the results of the feasibility study of the rationally-designed genetically encoded library approach.

LAY SUMMARY

In order for our bodies to grow and to replace lost cells, our cells must divide in a tightly controlled and orchestrated manner. To achieve this, the chromosomes that contain the cell's genetic information are first duplicated then segregated into two identical cells, termed daughter cells, which are then separated. It is imperative that each daughter cell receives one, and only one, copy of each of the duplicated chromosomes. Failure to maintain this fidelity can lead to cell death, birth defects and cancer.

The chromosomal passenger complex performs several essential functions during the process of cell division. In particular, the complex helps ensure that each of the two daughter cells contains an exact replica of the genetic information present in the original cell. The complex also serves key functions in cytokinesis – the process by which the two daughter cells become irreversibly separated. The chromosomal passenger complex exerts its functions by modifying other proteins via its Aurora B subunit. The other parts of the complex serve to activate Aurora B and direct its activity to where it is required.

Inhibiting Aurora B through different approaches is a useful strategy for understanding its function. This is also being pursued as an anti-cancer strategy as loss of Aurora B function preferentially kills cancerous cells, which are more sensitive than healthy cells to errors in cell division. Drugs designed to inhibit Aurora B all target its active site – the region through which it modifies other proteins. However, blocking key interactions between Aurora B and other chromosomal passenger complex members had been proposed as an alternative means of abrogating chromosomal passenger complex function by interfering with Aurora B activity and/or localisation.

To study the effects of blocking interactions within the chromosomal passenger complex, I devised an assay monitoring cell division in cancer cells. Using this assay, I demonstrate that disturbing the binding between Aurora B and inner centromere protein (INCENP) in cancer cells has a significant negative impact on cell division. Furthermore, I confirm the importance mimicking key features within the interaction for blocking it.

Further to validating the effects of blocking the interaction between Aurora B and INCENP, I wanted to determine whether it could be teased apart with small compounds, as this would suggest that it might be amenable to being inhibited with drugs. To test this, I created a library of small fragments resembling key features within the Aurora B and INCENP interaction. Although some fragments modestly affected cell division, the evidence gathered is too preliminary to reach conclusions about the potential of the interaction as a drug target.

CHAPTER 1: INTRODUCTION

1.1 PROTEIN-PROTEIN INTERACTIONS & PEPTIDES

Protein-protein interactions (PPIs) are viewed as promising, if technically challenging, drug targets (reviewed by (Sillerud and Larson, 2005; Wells and McClendon, 2007; Zutshi et al., 1998)). Such surfaces are typically more divergent than the enzyme active sites that have long been favoured in drug development campaigns. Inhibitors of active sites such as substrate analogues are often plagued by off-target effects due to the modular nature of, and high degree of conservation amongst, such surfaces. However, the large size of PPI surfaces relative to compounds with drug-like properties has discouraged efforts to develop inhibitors targeting these. The flat conformation of such contact surfaces also presents fewer binding pockets for small molecules. Another limiting factor has been the relative scarcity of suitable templates on which to base analogues with drug-like properties. A notable success in this area was the development human double minute 2 (HDM2, a.k.a. MDM2) binders based on information gained from structural and mutagenesis studies of a p53-derived peptide (reviewed by (Wells and McClendon, 2007). The use of peptide or miniprotein mimics to disrupt oligomerization interfaces has also yielded promising leads for antagonists of receptors and viral enzymes (reviewed by (Zutshi et al., 1998) and (Cochran, 2000).

TARGETING PROTEIN-PROTEIN INTERACTIONS WITH PEPTIDES

Peptide mimics with primary amino acid sequence homology to either interacting partner have an extensive track record as competitive inhibitors of protein-protein and protein-substrate interactions in *in vitro* assays. Short peptides that mimic binding motifs, dimerization interfaces and substrates

are routinely used in *in vitro* assays. Attempts to use such peptides against intracellular PPIs *in vivo* have been hampered by difficulties in delivering these into cells, and their poor stability therein.

DELIVERING PEPTIDES INTO CELLS

Attempts to assess the activities of short peptides in live cells have thus far been limited to extra-cellular applications or relied on bulky, and potentially interfering, tags or scaffolds – these are exemplified by cell penetrating peptides (CPPs) and peptide aptamers, respectively. Table 1.1 summarizes this and other notable applications of short peptide mimics as inhibitors and, when applicable, their method of delivery. For example, Yuan and colleagues fused a 19 residue long fragment of the polo kinase 1 (Plk1) polo box to a 16-mer antennapedia homeodomain, a cell penetrating peptide that acted as a carrier to translocate the polo box peptide into cells (Yuan et al., 2002). Soluble polo box was hypothesized to compete against endogenous Plk1 for substrate binding. Treatment with the polo box-antennapedia fusion peptide caused a G2/M arrest and induced apoptosis in multiple cell types compared to treatment with the 16-mer antennapedia domain alone. Point mutations to the polo box portion of the fusion peptide quelled its dominant negative effects.

Mimic	Target	Delivery	Notes	Ref
NFAT	Calcineurin	N/A	EA	(Aramburu et al., 1999; Roehrl et al., 2004)
HIV-1 Tat	TAR (RNA)	CPP*	ChemCP	(Lalonde et al., 2011)
$P21^{WAF}$	PCNA	N/A	IV	(Warbrick et al., 1995)
Survivin	Hsp90	CPP		(Plescia et al., 2005)
HIV-1 gp160	(self)	N/A	IV; EA	(Wild et al., 1992)
ASFV 54	dynein light chain	CPP		(Hernáez et al., 2010)
MEK-1	ERK	CPP		(Kelemen et al., 2002)
Plk1 polo-box	Multiple	CPP		(Yuan et al., 2002)
Pbs2p	Sho1p SH3 domain	aptamer	Y2H	(Woodman et al., 2005)

Table 1.1 – Selected examples of peptide inhibitors with primary amino acid sequence homology (CPP: cell penetrating peptide; ChemCP: chemically synthesized cyclic peptide; EA: extracellular application; IV: *in vitro* application; N/A: non applicable because of non-intracellular application, see 'notes' for further explanation; Y2H: yeast 2 hybrid). * HIV-1 Tat contains a well-characterized cell penetrating peptide sequence, its mimic retains properties known to aid cell permeability – namely: it is arginine-rich

PEPTIDE APTAMERS

Aptamers are a means of presenting either short peptide or nucleic acid epitopes *in vivo*. These genetically-encoded constructs rely on a constant nucleic acid or protein scaffold to stabilize and constrain a variable loop containing the epitope (reviewed by Hoppe-Seyler et al., 2004), a commonly used scaffold for peptide aptamers is that of *E. coli* thioredoxin – a 109 residue protein (Colas et al., 1996). As they are genetically encoded, aptamers lend themselves to the generation of large combinatorial libraries, which are typically screened against yeast two-hybrid reporters. This approach is usually the favoured application of peptide aptamers; they are seldom used to generate discrete rationally designed peptide mimics.

CELL-PENETRATING PEPTIDES

As non-covalently bound carriers or as fusion proteins, CPPs can translocate cargo such as proteins and nucleic acids across the plasma membrane independently of the endocytic pathway (reviewed by Järver and Langel, 2004; Lindgren et al., 2000). Two major CPP subtypes have been identified thus far, consisting of either ampiphatic or positively charged stretches of residues. An example of the latter category is the transactivation domain of the HIV-1 protein Tat (Frankel and Pabo, 1988) – the most predominantly used CPP for peptide delivery. First identified in 1988, the minimal region required for translocation was eventually narrowed down to an arginine-rich stretch of approximately a dozen residues (Vivès et al., 1997). The use CPP fusions in complex systems has been hampered by poor cell-type specificity in uptake (Järver and Langel, 2004). Toxicity is another drawback: some CPP types have been shown to be cytotoxic and it has also been reported that, even though the minimal Tat CPP is non-toxic, larger Tat fusion peptides exhibit a general length-dependent cytotoxicity (Cardozo et al., 2007).

SPLIT INTEIN CIRCULAR LIGATION OF PROTEINS AND PEPTIDES

In contrast to the approaches discussed above, the split intein circular ligation of proteins and peptides (SICLOPPS) system allows for the intracellular production of genetically encoded cyclic peptides (CPs) without recourse to tags. Developed by Stephen Benkovic and colleagues in the late nineties, the technique relies on the naturally split intein from the *Synechocystis sp.* PCC6803 (*Ssp*) DNA polymerase III subunit *DnaE* (Scott et al., 1999). As illustrated in panel A of Figure 1.1.1, the two halves of the *Ssp DnaE* helicase are encoded by two separate reading frames, in different orientations, separated by 745kb on the *Ssp* genome (Wu et al., 1998).

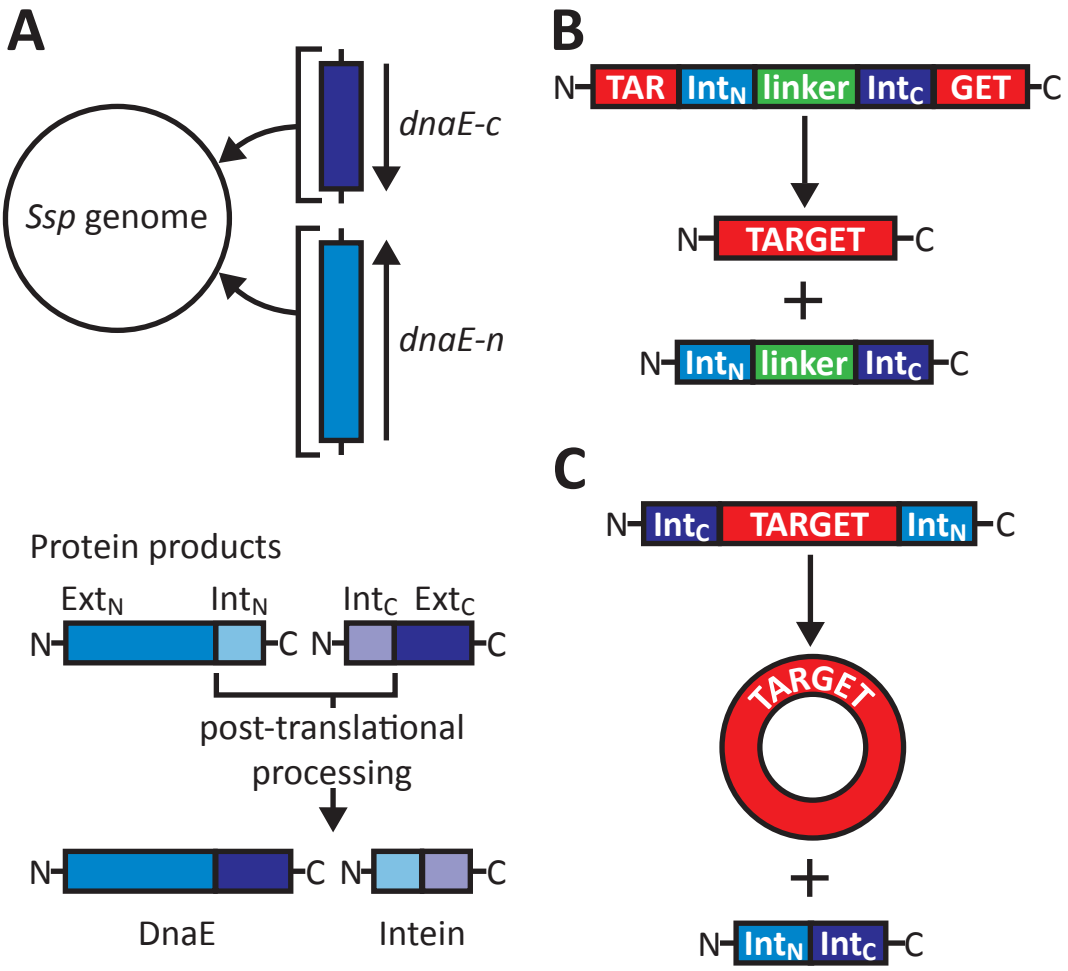


Figure 1.1.1 - Principle of the SICLOPPS method. (A) The N- and C- terminal parts of the Ssp DnaE split intein originate from two separate open reading frames. These interact in trans to catalise the post-translational splicing of the DnaE N- and C- terminal peptides to form the mature protein (after Wu et al., 1998). (B) Spit inteins can also ligate proteins and peptides in cis. If the native intein order is maintained, sequences which are proximal and distal to the respective halves of the split intein are joined upon intein processing. (C) In contrast to this, reversing the order of the intein halves leads to the head-to-tail circularisation of the intervening sequence (after Abel-Santos et al., 2003). Legend: ext: extein; Int: intein; N: amino; C: carboxy.

RATIONALE

In their endogenous context, a split intein catalyzes the post-translational protein splicing of N- and C-terminal exteins in *trans* whilst simultaneously excising itself from the protein product. In the instance of the *Ssp DnaE* split intein, processing results in seamless ligation of the two gene translation products thereby reconstituting the active *DnaE* helicase. Split inteins can also act in *cis* and catalyse the head-to-tail circularization of the intervening protein sequence when their order is reversed (Figure 1.1.1, panel B and C, respectively) – therein lies the crux of the SICLOPPS technique.

One can generate a desired CP by encoding its sequence in an intervening stretch of residues between the C- and N-terminal inteins. Polypeptides reported to have been circularized using this approach range in length from whole proteins to peptides as small as 4 residues (Scott et al., 2001; Scott et al., 1999). By generating the extein region of the SICLOPPS construct (i.e. that encoding the fragment which will be excised and circularized) using randomized oligos, one can generate very large combinatorial libraries encoding highly diverse CPs (Scott et al., 2001). Typical applications of the SICLOPPS methodology have involved screening such combinatorial libraries.

ADVANTAGES

CPs possess numerous advantages over their linear counterparts (reviewed in Craik, 2006; Trabi and Craik, 2002). Due to the lack of termini accessible to exopeptidases, head-to-tail circularization of peptides confers greater stability (for example, see Szewczuk et al., 1992). Peptide circularization is a popular approach for enhancing a peptide's affinity. This is thought to be achieved by limiting a ligand's conformational flexibility thereby reducing changes in entropy upon binding (Khan et al., 1998). In

practice however, it is possible for binding entropy to be enhanced, reduced or unaltered between cyclic peptides and linear analogues (Udugamasooriya et al., 2008).

Further advantages of the SICLOPPS technique lie in its ease of delivery and minimal requirements for extraneous sequence. Firstly, in common with aptamers, these constructs are genetically encoded. This circumvents the issues surrounding peptide cell-permeability. Furthermore, the protein splicing is autocatalytic, requiring no accessory factors. The technique is, therefore, amenable to any cell types that can be readily transfected. Lastly, the only absolute extein sequence requirement is a nucleophillic residue in position +1 of the circularized fragment (Wu et al., 1998). This represents a vast improvement on the length of accessory residues required for either aptamer- or CPP-mediated peptide delivery.

APPLICATIONS

The sheer size and complexity of combinatorial libraries raises both tantalizing prospects and logistical headaches. Of particular concern are the limitations imposed on screen strategy by the nature of such libraries, and difficulties in target identification. Combinatorial genetically-encoded CP libraries must be produced in pooled formats (Abel-Santos et al., 2003). Screening them therefore requires either strong positive selection or resolution by deep sequencing. Screens published to date have relied on the former – either by lethality suppression or reverse two-hybrid approaches (Horswill et al., 2004; Kritzer et al., 2009). This strategy serves two purposes: to enrich for positive clones; and to permit the identification of the CP giving rise to the desired phenotypic outcome. The latter is achieved by sequencing the library construct present in positive clones.

However, cell survival is not theoretically essential for recovering genetically encoded hits screen *in vivo*. The potential of dropout screens comparing population-wide CP distribution before and after selection has been investigated in *S. cerevisiae* (R. Almeida & C. Arndt, personal communication). In that instance, the variable sequence of the CP serves a purpose analogous to a yeast barcode and screen hits are the sequences present in the population prior to, but not after, phenotypic selection.

Despite the suitability of intein-mediated peptide circularisation to a wide variety of cell types, screens of combinatorial CP libraries have almost exclusively been carried out in prokaryotes or lower eukaryotes. To date, a single attempt to screen such a CP library in human cells – generated using an artificially-split *Ssp DnaB* intein – has been published: in 2002, Kinsella and colleagues tested a relatively low complexity ($1.6x10^5$ members) cyclic pentapeptide library in human B cells to find inhibitors of interleukin-4 (IL-4) signalling (Kinsella et al., 2002). In this study, researchers engineered a reporter containing the promoter of the germ line epsilon gene that, if expressed, would sensitize cells to subsequent exposure to diphtheria toxin. In B cells, the epsilon gene promoter is induced by IL-4 signaling. This screen relied on lethality suppression read out: hits were isolated by finding diphtheria toxin-resistant cells under conditions where IL-4 signaling should be active. The success of this study was deemed limited as the 13 hits identified through screening did not show higher activity than existing inhibitors and their mechanism of action could not be ascertained.

Compared to cell based screens employing reporters such as reverse two hybrid systems, phenotypic screens do not necessarily require prior target identification (reviewed in Butcher, 2005). In theory, this permits the identification of novel players and pathways giving rise to the outcome of

interest. However, whilst phenotype-based screens of combinatorial CP libraries have yielded hits, it is disappointing that no interacting partners for these have thus far been identified despite attempts to do so (Kinsella et al., 2002; Kritzer et al., 2009). For example, Kritzer and colleagues prepared affinity reagents based on their most potent CP hits but were unable to pull down binding partners. Whilst preliminary structure-activity relationship studies can still be carried out using a phenotypic assay, failure to identify interacting partners, and by extension probable targets, can hamper further lead optimization and constitutes a translation roadblock.

1.2 CELL DIVISION AND CANCER

In multi-cellular organisms, controlled cell division is required to meet the demands of development, growth and tissue turnover. As the cell cycle progresses, the genome is first replicated then partitioned into two daughter cells. Further to the genome, other cellular structures such as mitochondria, the endoplasmic reticulum and Golgi bodies must also be replicated and partitioned. Due to their relevance to this project, I will principally focus here on the dynamics of genomic material throughout the vertebrate cell cycle.

An overview of the cell cycle is presented in Figure 1.2.1. Cells first undergo a growth phase (gap 1 or G_1), followed by a synthesis phase (S; in which DNA synthesis takes place) then a second growth phase (gap 2 or G_2). Together, this series of steps are known as interphase. The newly replicated genome is subsequently partitioned in mitosis (M) – a dynamic and tightly orchestrated process (see below). Many cells do not actively divide and are in a quiescent state (known as gap 0 or G_0); some of them can re-enter the cell cycle given appropriate external stimulus whereas others cannot. Progress through the cell cycle is tightly controlled and subject to a number of checkpoints.

MITOSIS & CYTOKINESIS – AN OVERVIEW.

Mitosis is the process by which the replicated genome is faithfully segregated into two daughter nuclei and cytokinesis is that by which two daughter cells are formed. These are highly dynamic processes involving drastic changes in chromosome, microtubules and membranes (Robbins and Gonatas, 1964). As cells enter mitosis, the diffuse and accessible genomic material present in the interphase nucleus undergoes drastic morphological and spatial rearrangements. Mitotic phases and cytokinesis were observed and

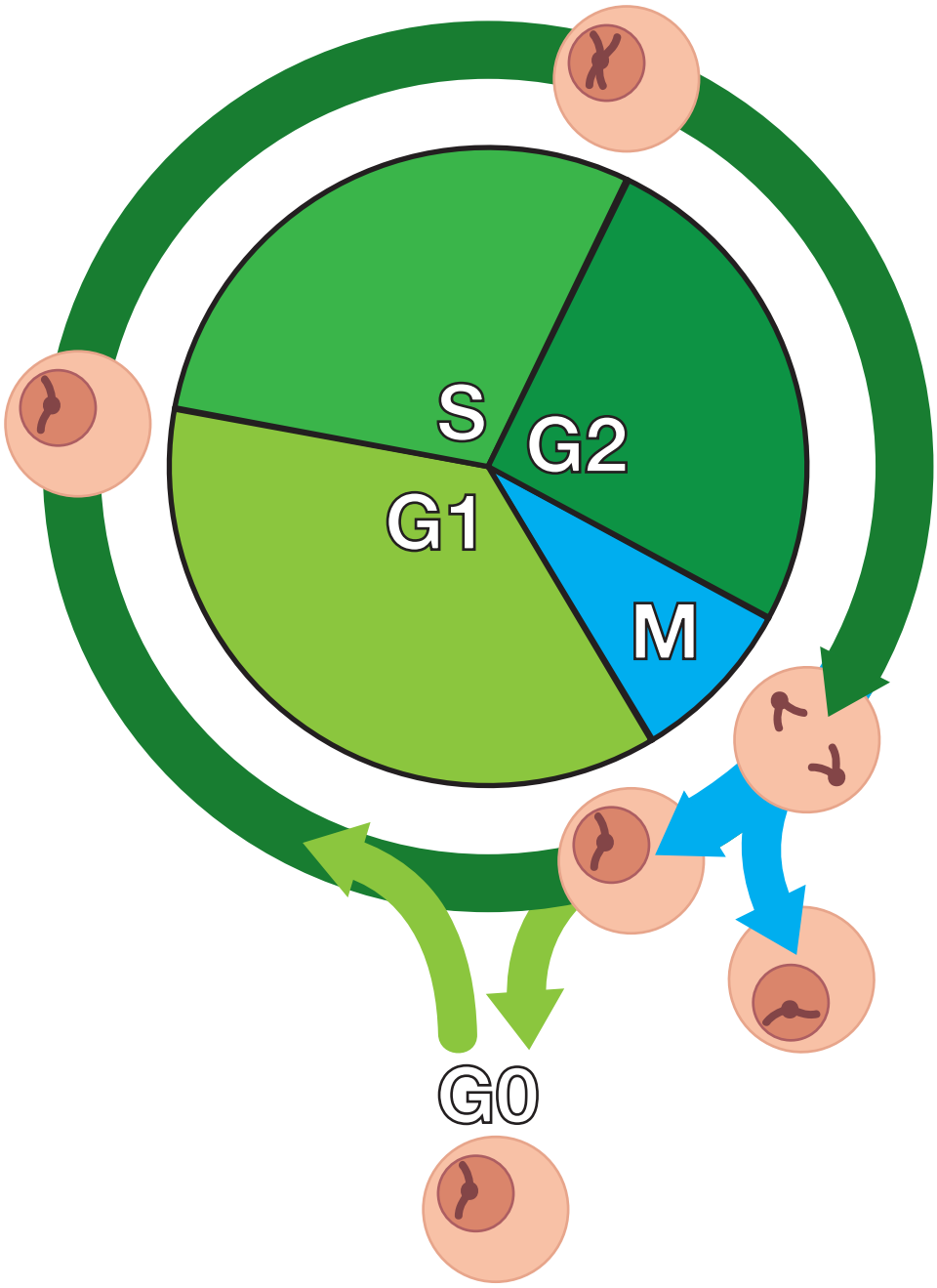


Figure 1.2.1 - The cell cycle. Following mitotic exit, newly created daughter cells enter interphase (green). Interphase constitutes the majority of the cell cycle and is interupted by S phase, during which the genome is replicated. The duplicated genome is subsequently partitioned into two daughter nuclei through mitosis (blue) before a cytoplasmic division creates two daughter cells - this latter process is known as cytokinesis. Following division, cells can also exit the cell cycle and become quiescent (G0: quiescence; G1: gap1; G2: Gap 2; M: mitosis; S: synthesis).

described by early cytologists based on microscopy observations. The hallmarks of each phase are highlighted below; for an illustrative overview of these, please consult Figure 1.2.2.

1. ***Prophase***

Centrosome separation occurs and individual centrioles migrate to opposite poles of the cell – these will form the focused poles from which spindle microtubules emanate. Within a still intact nuclear envelope, chromosomes condense and individual sister chromatid pairs become visible.

2. ***Prometaphase***

Nuclear envelope breakdown occurs allowing for spindle microtubules to enter the chromosome-containing compartment. Chromosomes associate with said microtubules through both lateral and end-on attachments and start to congress towards the center of the mitotic spindle.

3. ***Metaphase***

Metaphase is a transient but key mitotic phase: the point at which all chromosomes have congressed to the metaphase plate (i.e. the spindle midzone) and become bi-oriented. Bipolar attachment of sister chromatids generates tension across their kinetochores and leads to the cessation of local production of a 'wait anaphase' signal. Once all sister chromatid pairs are bi-oriented, the lack of unattached kinetochores represses the inhibition of the anaphase promoting complex, leading to the proteolytic degradation of cyclin B and other substrates (this is elaborated in 'The mitotic checkpoint' section below.)

4. ***Anaphase***

The link between sister chromatids is broken as centromeric cohesion is cleaved. Spindle forces pull apart the sister chromatids causing these to segregate towards opposite poles.

5. ***Telophase***

The spindle begins to disassemble, the nuclear envelope reforms around each daughter nucleus and chromosomes begin to decondense within these. This is the final phase of nuclear division.

6. ***Cytokinesis***

The separation of the two daughter cells begins during telophase. Driven by an actomyosin contractile ring, the equatorial cortex (i.e. the cell membrane situated over the spindle midzone) begins to ingress, forming the cleavage furrow. This cleavage furrow then continues to ingress until it reaches the midbody (i.e. the transient structure derived from the spindle midzone, which is necessary for cellular division). The thin intracellular bridge between the two daughter cells is then resolved through abscission. It should be noted that although cytokinesis typically follows mitosis, the two processes are distinct and can be uncoupled.

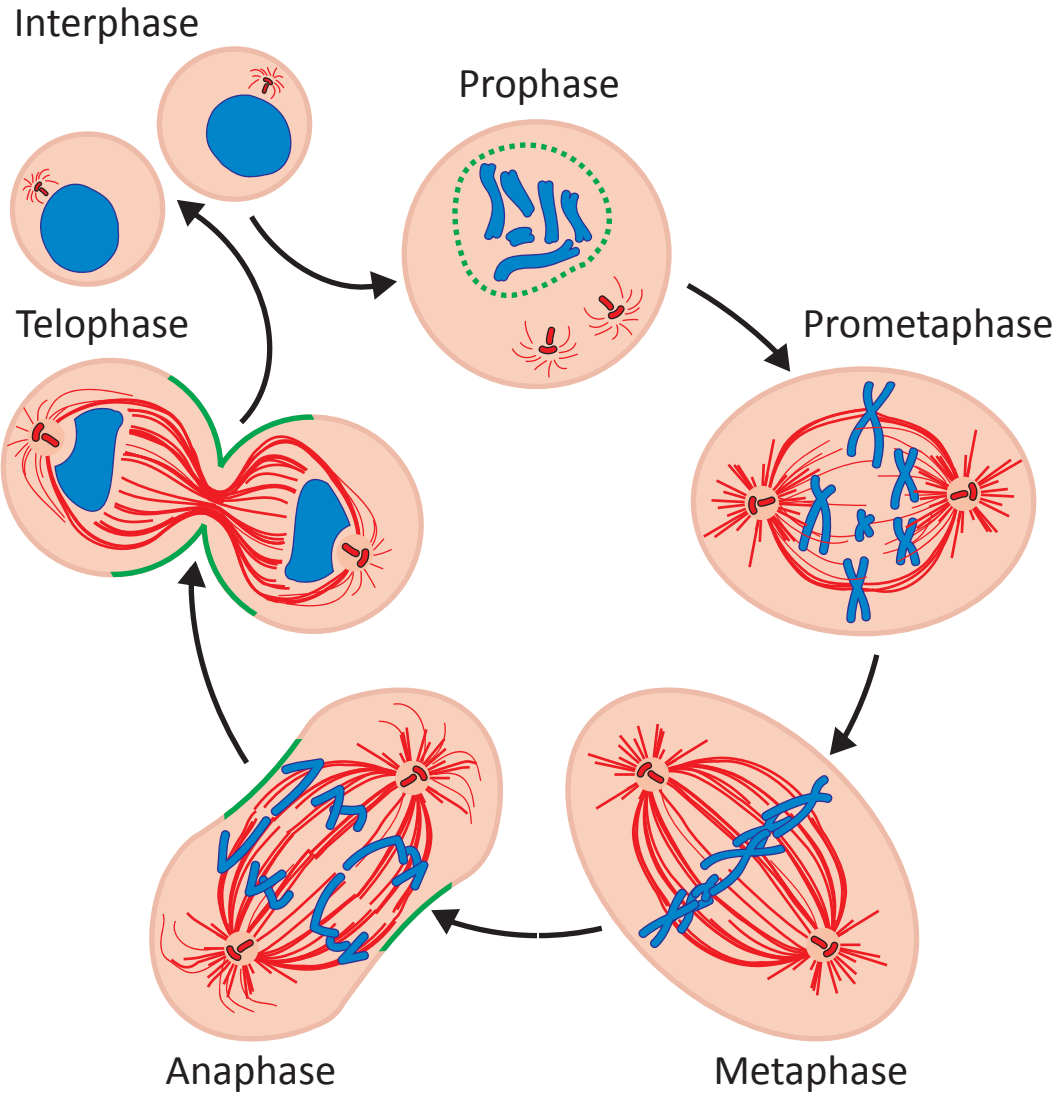


Figure 1.2.2 - The stages of mitosis illustrated with respects to chromosome (blue), spindle (red) and membrane (green) dynamics (after Ruchaud et al., 2007).

THE MITOTIC CHECKPOINT

Failure to maintain genetic fidelity in mitosis has deleterious circumstances and can result in cell death or genetic aberrations that have the potential to drive neoplastic transformations (e.g. by loss of tumour suppressor genes). Like S phase, progression through mitosis is controlled by checkpoints to ensure genomic integrity.

Anaphase onset and the separation of sister chromatids must be prevented until all sisters are bi-oriented. This is achieved by the inhibition of the anaphase promoting complex/cyclosome (APC/C) by a mitotic checkpoint complex (MCC) that is generated at unattached kinetochores (Kulukian et al., 2009; Sudakin et al., 2001) – the specific role of Aurora B and the CPC in this process will be elaborated upon in section 1.3. As chromomes achieve biorientation, MCC production ceases and the APC/C becomes activated. The APC/C is an E3 ubiquitin ligase; its activation triggers the ubiquitination and destruction of substrates containing an epitope known as the destruction box (D box) (King et al., 1995). Substrates include cyclin B and securin, an inhibitor of the protease separase whose substrates include cohesin (Hagting et al., 2002; King et al., 1995). Cohesin cleavage resolves pericentromeric cohesion and the linkage between sister chromatids, which can then segregate to opposite poles (Uhlmann et al., 1999).

CANCER THERAPY – A PRIMER

Cancer is usually a disease characterized by unfettered cell division driven by genomic lesions. Cancerous tissues typically proliferate rapidly and thus possess elevated mitotic indexes relative to most healthy tissues. Furthermore, the compromised checkpoints that permit deregulated proliferation can lead to cells proceeding through the cell cycle despite the presence of faults, which can lead to cell death. These features have been

exploited to preferentially target cancerous cells with cytotoxic therapies. With few relatively recent exceptions, cancer therapies have either targeted DNA metabolism or mitosis (the latter is discussed in greater detail in the next section.)

Although the consequences of treatment with DNA damaging agents are suffered most acutely by actively cycling cells, the genomic integrity of all cells – including quiescent cells – can be affected. Further to deleterious side effects due to the death of healthy but rapidly proliferating cells (e.g. neutropenia), treatment with DNA damaging agents such as radiation risks giving rise to an increased frequency of subsequent neoplasms in the future (Boice et al., 1985; Brenner et al., 2000).

Further challenges in treating cancer stem from genomic instability – a hallmark of cancer that generates diversity within the cancerous cell population. This diversity creates significant therapeutic challenges as treatment drives resistance by positively selecting for drug-resistant sub populations. Mechanisms of resistance include the overexpression of drug efflux pumps and the acquisition of mutations to the drug target site (Gottesman, 2002). As it is unspecific, the former confers multi-drug resistance. A well-characterized instance of the latter mode of resistance is the acquisition of so-called 'gate-keeper' mutations conferring resistance to imanitib – a small molecule inhibitor of the Bcr-Abl oncogenic gene fusion product (Shah et al., 2002). Therapeutically speaking, like human immunodeficiency virus (HIV) infection, cancer is considered a 'moving target'.

As in the HIV field, combination therapy is used in the treatment of cancer. Furthermore, our growing understanding of the genetic lesions capable of driving the proliferation of transformed cells is creating new

avenues for the development of novel therapeutic agents. Additionally to generic approaches targeting key phases of the cell cycle, substantial investments are being made into agents exploiting tumour-specific drivers and weaknesses along with companion diagnostics for use in personalized medicine treatment approaches.

ANTI-MITOTICS AS CANCER THERAPEUTICS

As previously highlighted, the potential for collateral damage from DNA damaging agents extends to quiescent cells. The rationale for interfering with mitotic progression as anticancer therapeutic strategy, in contrast to DNA damaging agents, was to limit such potential damage to actively cycling cells. As their name implies, anti-mitotics are agents disrupting progression or completion of mitosis. The first such agent introduced into the clinic were vinca alkaloids – small molecule microtubule destablising agents.

Microtubule-targeting agents have proven effective in the clinic. Early anti-mitotic agents now in clinical use include microtubule poisons such as vinca alkaloids and taxanes, which respectively promote microtubule destabilization and stabilization at high concentrations and a concomitant block in mitosis (Jordan and Wilson, 2004). Exposure of cultured cells and tumour xenografts to microtubule poisons results in mitotic arrest, slippage and eventual cell death. Although increases in tumour mitotic indexes can be observed in most patients treated with paclitaxel, these does not necessarily correlate with apoptotic indexes or overall tumour treatment response (Symmans et al., 2000). However, as microtubules are also essential components of non-cycling cells, microtubule-targeted agents have unintended and potentially limiting side effects. Peripheral neuropathy is a particular concern and may be linked to disrupted axonal transport, which requires dynamic microtubules (Lee and Swain, 2006).

The therapeutic success and unintended side effects of microtubule poisons have motivated the search for drug targets whose activity is wholly restricted to mitosis. Mitotic kinesin inhibitors, which disrupt spindle formation, have entered clinical trials (Huszar et al., 2009). Inhibitors of mitotic kinases including cyclin-dependent kinases, polo and Aurora kinases have also been developed and significant efforts have been invested into their clinical development (Keen and Taylor, 2004; Lens et al., 2010).

It is important to caution that, although efficacious in pre-clinical models such as culture cancer cells and rodent xenographs, the performance of these newer anti-mitotic compounds in clinical trials has been disappointing thus far. These compounds are no better tolerated than other chemotherapeutic agents and disease response has been modest at best. Poor clinical efficacy relative to microtubule targeting agents has recently lead some to suggest that the mode of action of microtubule poisons may transcend their effects in mitotic cells and encompass other essential processes such as vesicle trafficking and angiogenesis (Komlodi-Pasztor et al., 2011; Mitchison, 2012). Furthermore, as summarized within those opinion pieces, the doubling times – thus the portion of cells in mitosis – within tumours can be much lower than those in cell lines and animal xenograft models. These results put into question the rationale for the specific targeting of mitotic activities and suggest that the action of effective anti-tumour therapeutic strategies, including those affecting DNA replication, may extend beyond specific cell cycle phases.

1.3 AURORA B AND THE CHROMOSOMAL PASSENGER COMPLEX

The CPC orchestrates a number of key events during mitosis and shows dynamic, and highly characteristic, spatio-temporal localization throughout this process (reviewed in Ruchaud et al., 2007) – its composition and localization are outlined in figures 1.3.1 and 1.3.2, respectively. The localization of the CPC was first described by Cooke and colleagues using anti-INCENP antibodies (Cooke et al., 1987). At the onset of mitosis, the CPC is present both on chromosome arms and at centromeres, promoting mitotic chromosome structure via the phosphorylation of substrates including histones and chromosomal scaffold proteins. CPC localisation gradually becomes increasingly centromeric as mitosis progresses towards metaphase, wherein it promotes chromosome bi-orientation and ensures integrity of the spindle checkpoint until bi-orientation is achieved. The complex then translocates to the equatorial cortex and spindle midzone at anaphase and helps stabilize the central spindle. The complex eventually accumulates in the midbody in cytokinesis wherein it regulates contractile ring formation and abscission. The roles of the CPC are explored in greater detail later in this chapter.

Further to Aurora B, the CPC contains a further three non-enzymatic subunits: the inner centromere protein (INCENP) (Cooke et al., 1987), survivin (Bolton et al., 2002) and borealin/Dasra-B (hereafter referred to as borealin; (Gassmann et al., 2004; Sampath et al., 2004). Whilst the latter two members function as targeting subunits, INCENP serves dual roles within the complex: both as a scaffolding subunit, and as an activator of Aurora B. CPC composition and function are highly conserved throughout eukaryotes but the following literature review will focus specifically on the CPC in humans.

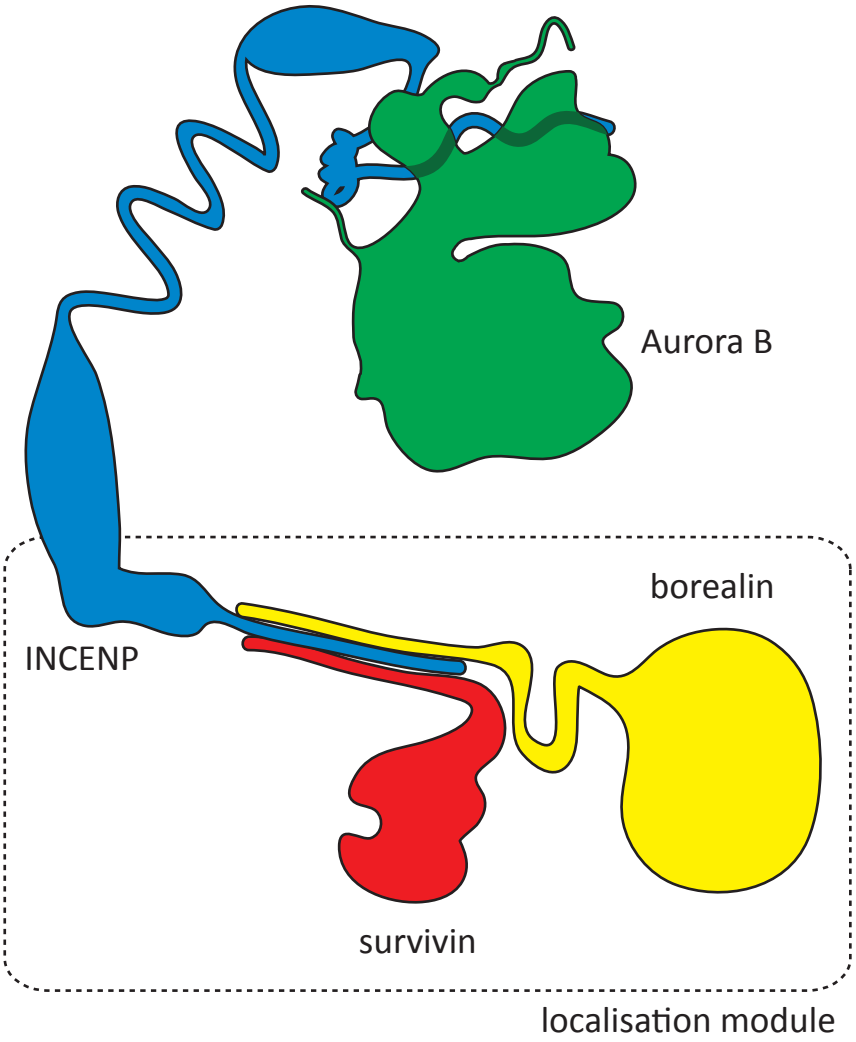


Figure 1.3.1 - The composition of the chromosomal passenger complex. The kinase Aurora B is responsible for the complexe's catalytic activity, and borealin, survivin and INCENP form a localisation module required to direct the CPC's activity. The three members of the localisation module interact via a triple-helix bundle (after Carmena et al., 2012).

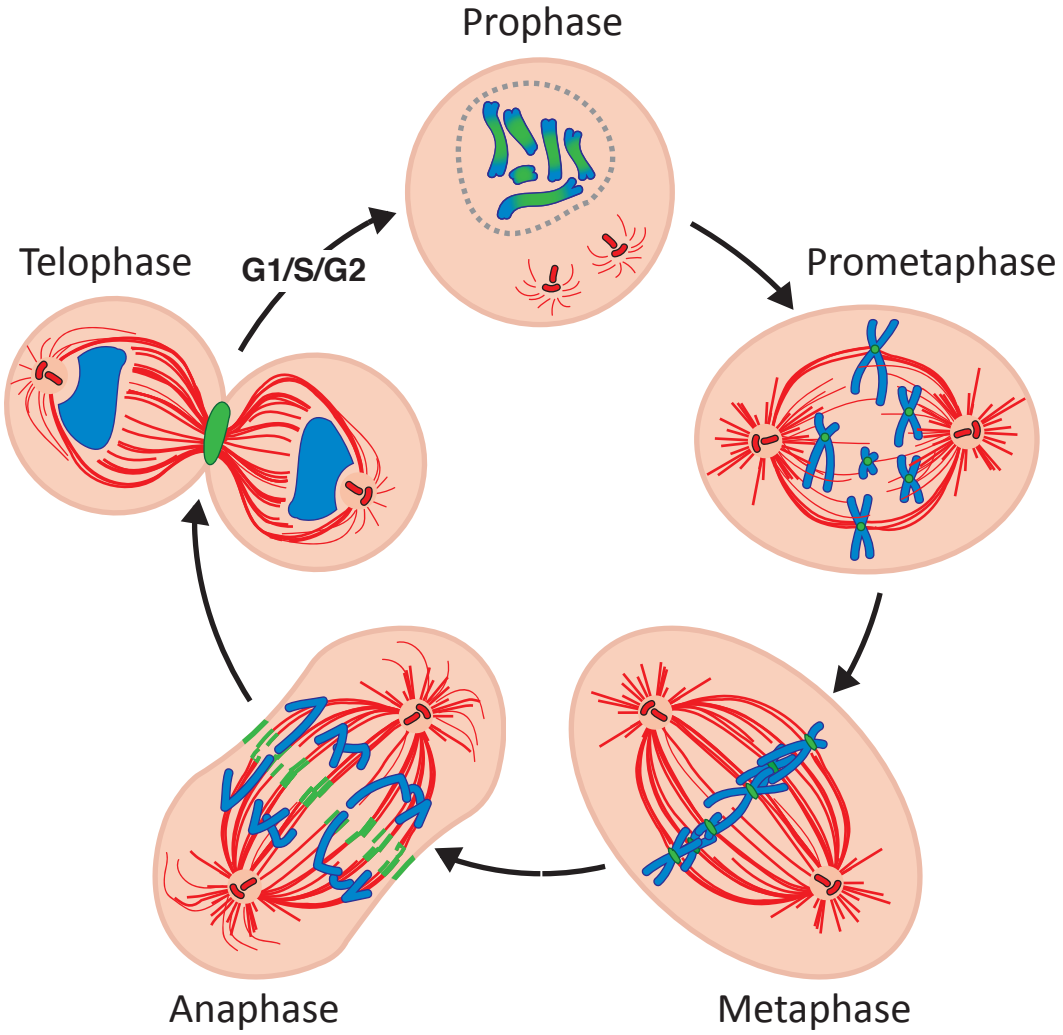


Figure 1.3.2 - Localisation of the chromosomal passenger complex during mitosis. The CPC (green) is first localised to chromosomes arms (blue), then to the centromeres until metaphase. Upon anaphase onset, the CPC relocalises to the spindle (red) and eventually becomes concentrated into the midbody (after Ruchaud et al., 2007).

Aurora B is a member of a conserved family of serine/threonine kinases. A further two paralogues are present in vertebrates: Aurora A and C. Aurora A is widely expressed and plays a role in centrosome maturation and spindle organization (Glover et al., 1995) whereas Aurora C, whose expression is restricted to the germ line and early embryos, is essential for male meiosis and early embryogenesis (Avo Santos et al., 2011; Balboula and Schindler, 2014; Fernández-Miranda et al., 2011; Kimmins et al., 2007). First identified in *Drosophila* by Glover and colleagues, Aurora A was the first member of the Aurora kinase family to be discovered.

Despite their divergent functions, there is a high degree of conservation amongst the Aurora kinases. The respective localizations and substrate specificities of the Auroras are directed via their cognate binding partners (reviewed in Carmena et al., 2009). The CPC serves this adaptor function for both Aurora B and C, which both bind to INCENP via its INbox domain (Adams et al., 2001; Li et al., 2004). Aurora A is known to interact with a number of adaptors, the best characterized of which is the microtubule-binding protein TPX2 (Kufer et al., 2002).

That substrate specificity of Aurora kinases is achieved through localisation has been highlighted in a number of studies. Most notably, a single-residue mutation in the TPX2-binding domain of Aurora A to the analogous residue in Aurora B leads to INCENP binding, Aurora B-like localization, and can rescue the loss of Aurora B (Hans et al., 2009). Ectopic expression of Aurora C can also rescue Aurora B loss of function (Sasai et al., 2004). Conversely, expression of a kinase-dead Aurora C mutant interferes with CPC function reiterating effects of kinase-dead Aurora B expression (Chen et al., 2005; Girdler et al., 2006; Sasai et al., 2004).

ACCESSORY CPC SUBUNITS: INCENP, SURVIVIN & BOREALIN

As outlined earlier, the CPC consists of Aurora B, INCENP, survivin and borealin. The stoichiometry of the full complex is still to be elucidated but survivin and N-terminal fragments of INCENP (1-58) and borealin (10-109) form a 1:1:1 complex (Jeyaprakash et al., 2007) and Aurora B and INCENP form either a 1:1 or a 2:2 complex, the latter of which through an activation loop exchange between two Aurora B subunits (Sessa et al., 2005; Elkins, 2012). Purified Borealin forms a homodimer (Bourhis et al., 2009), as does purified survivin (Chantalat et al., 2000; Verdecia et al., 2000). Together however, purified borealin and survivin form a 1:1 complex (Bourhis et al., 2007). A further protein, telophase disk 60 protein (TD60), shows passenger-like localization but is not stably associated with the complex itself and will not be considered here (Martineau-Thuillier et al., 1998).

As illustrated in Figure 1.3.2, the CPC possesses activity (Aurora B) and localization (survivin and borealin) modules with INCENP providing an elongated flexible linker between the two. One current model proposes that INCENP acts as 'leash' through which Aurora B is flexibly tethered to specific sites recognized via the tripartite localization domain (Santaguida and Musacchio, 2009). This would permit the kinase limited but flexible reach to phosphorylate substrates. The localization module is formed via a tripartite interaction between the INCENP N-terminus, borealin and the C-terminus or survivin. All three domains adopt a helical conformation and interact with each other to form a triple helix bundle (Jeyaprakash et al., 2007). The C-terminus of INCENP crowns the N-terminal lobe of Aurora B (Sessa et al., 2005). This interaction not only links Aurora B to the localization domain but, most importantly, activates the kinase (see 'Aurora B activation' section).

INCENP

First described in by Cooke and colleagues 1987, INCENP was the first CPC component to be identified (Cooke et al., 1987). INCENP was discovered via a chromosome scaffold monoclonal antibody screen by virtue of its highly dynamic and stereotypical mitotic localization. The epitope was found to associate with the inner centromere up until metaphase (giving the protein its name), and then transfer to the spindle midzone and equatorial cortex upon anaphase onset – the hallmarks of what is now known as 'passenger-like' localization.

INCENP serves two primary functions: it acts as a scaffold for the CPC by linking its localization and activity modules via its highly conserved termini, and it activates Aurora B kinase. The latter function will be discussed in detail in the 'Aurora B activation' section. The INCENP N-terminus is bound by the two localization subunits of the CPC: survivin and borealin; its C-terminus contains the INbox, via which it binds and activates Aurora B. Distal to INCENP's N-terminal alpha helical survivin and borealin interaction domain is a heterochromatin protein 1 (HP1) interacting domain through which INCENP interacts with chromatin (Ainsztein et al., 1998). The intervening extended stretch of residues linking the two INCENP termini is less highly conserved and structured. This domain is also known to associate with microtubules (Wheatley et al., 2001) and is posited to offer flexibility and 'reach' for Aurora B.

BOREALIN

Also known as Dasra, borealin was identified simultaneously through two independent studies investigating proteins associated with mitotic chromosomes or the mitotic scaffold (Gassmann et al., 2004; Sampath et al., 2004). As indicated earlier, borealin participates in the formation of a triple

helix bundle, which it does via its N-terminus. The central region of borealin interacts with the endosomal sorting complexes required for transport III (ESCRT-III) complex – a cytosolic membrane remodeling complex required for abscission (see 'Roles of the chromosomal passenger complex – in cytokinesis' below.) As mentioned earlier, purified borealin forms a homodimer, which it does via its C-terminus. A Mps1 phosphorylation site within this dimerization domain is required for Aurora B activity in mitosis (Bourhis et al., 2009).

SURVIVIN

Survivin is an essential protein first identified as a member of the inhibitor of apoptosis (IAP) family by virtue of its zinc-coordinated N-terminal BIR (baculovirus IAP repeat) domain (Ambrosini et al., 1997). Although first presumed to possess an anti-apoptotic role by virtue of its conserved BIR domain, this domain binds to mitotic substrates with much greater affinity than to the pro-apoptotic Smac/DIABLO peptide (Du et al., 2012). Moreover, survivin does not bind to apoptotic protease caspase 3 (Banks et al., 2000; Li et al., 2008). Cells lacking survivin show increased chromosome alignment and cytokinesis defects, which are consistent with it possessing mitotic functions (Carvalho et al., 2003; Lens et al., 2003).

The C-terminus of survivin consists of a helical domain that interacts with itself to form a homodimer, or with borealin and the INCENP N-terminus to form the triplex helix bundle described earlier. The survivin N-terminal BIR domain is necessary for CPC localization to the centromere, but not the spindle, in vertebrates (see 'Chromosomal passenger complex localisation' section below). In human and chicken DT40 cells, survivin BIR mutants fail to recruit Aurora B in early mitosis but do not affect the kinase's post-anaphase

relocation to the spindle midzone and midbody (Lens et al., 2006; Yue et al., 2008).

AURORA B ACTIVATION

Activation of Aurora B is mediated by interactions between the N-terminal lobe of the kinase and the highly-conserved the INCENP C-terminus: the INbox (Adams et al., 2001; Adams et al., 2000; Kaitna et al., 2000). Full activation of Aurora B by the INCENP INbox occurs via a sequence of phosphorylation events (Bishop and Schumacher, 2002). A crystal structures of the two proteins has revealed that the INbox does not make contact with either the Aurora B catalytic cleft or its activation loop (Figure 1.3.3). Rather, it allosterically promotes the opening of the catalytic cleft and *trans* activation (Elkins et al., 2012; Sessa et al., 2005). Local enrichment of Aurora B and INCENP is also required for the kinase's activation, as demonstrated by the artificial clustering of INCENP-bound Aurora B using anti-INCENP antibodies (Kelly et al., 2007).

INbox binding to Aurora B generates the fully active CPC complex through a sequence of phosphorylation events. Upon INCENP binding, Aurora B phosphorylates a conserved threonine-serine-serine (TSS) motif located in the C-terminus of the INbox (Bishop and Schumacher, 2002; Honda et al., 2003). Phosphorylation of the TSS motif promotes phosphorylation of threonine 232 on the Aurora B activation loop by Aurora B in *trans* (Yasui et al., 2004). A recent xray crystallography study of the human Aurora B-INbox interaction detected a diametrically arranged structure in which Aurora B activation loops are exchanged between adjacent Aurora B-INbox dimers (Elkins et al., 2012). Although the biological significance of this arrangement is unclear, this evidence supports the *trans* activation model.

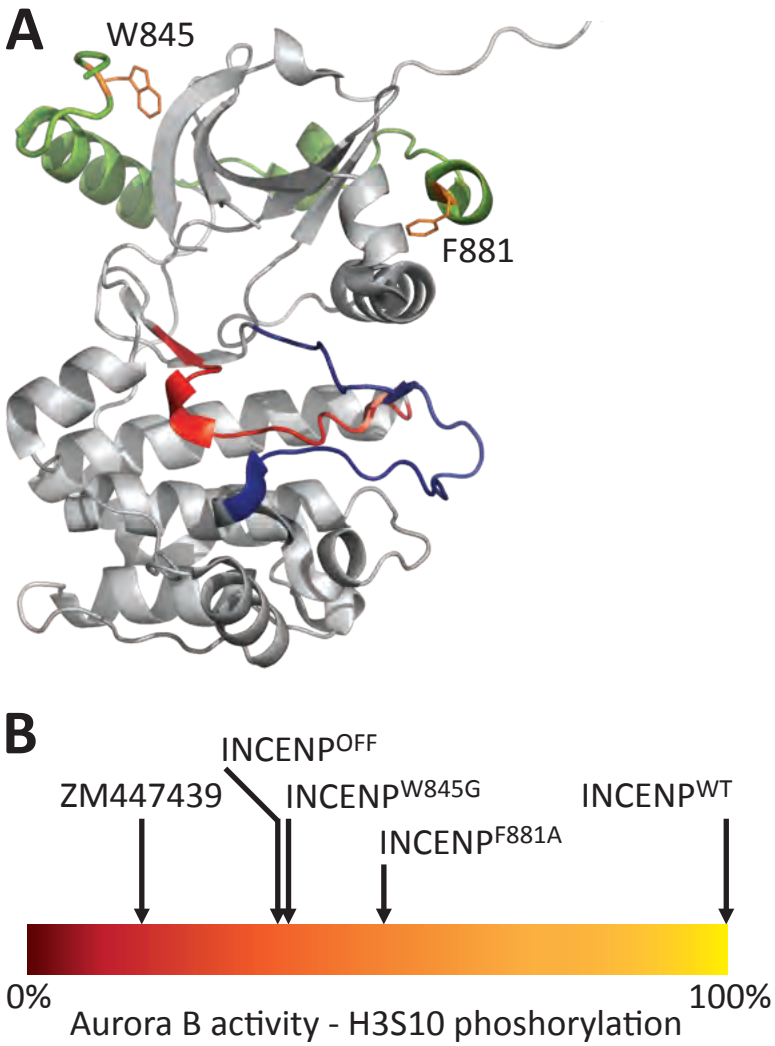


Figure 1.3.3 - Activation of Aurora B by the INCENP INbox. (A) The INCENP INbox (green) wraps around the N-terminal lobe of Aurora B (gray) and does not make contact with either the activation or catalytic loop (blue and red, respectively). Residues highlighted in orange are the two absolutely conserved residues that, when mutated, result in loss of Aurora B activation by the INbox (after Sessa et al., 2005; PDB:2BFY). (B) INCENP INbox mutants can only partially rescue Aurora B activity levels in chicken DT40 INCENP knockout cells (INCENP^{OFF}). Basal Aurora B activity present in the knockout is further diminished by treatment with the Aurora B inhibitor ZM447439 (after Xu et al., 2009). For consistency, residues are numbered according to their homologous coordinates in human INCENP.

Mutagenesis of both the TSS motif and other highly conserved residues within the INbox has yielded a series of CPC hypomorphs (Figure 1.3.3; Sessa et al., 2005; Xu et al., 2009; Xu et al., 2010). Previous work from members of our group has demonstrated that the chicken mutant of INCENP equivalent to human INCENP^{W845G} is incapable of binding and activating Aurora B, whereas the one equivalent to INCENP^{F881A} binds the kinase, but is activation-deficient (Xu et al., 2009). Neither of these mutants can rescue viability in the absence of endogenous INCENP. As indicated in Figure 1.3.3, in the absence of INCENP Aurora B possesses a detectable basal activity that is sensitive to treatment with active site inhibitors.

CHROMOSOMAL PASSENGER COMPLEX LOCALISATION

CPC localisation throughout mitosis is dynamic and tightly orchestrated. As outlined earlier, the CPC is first present both on chromosome arms and at centromeres in early mitosis. Its localisation gradually becomes restricted to the centromere as mitosis progresses towards metaphase. Upon anaphase onset, the complex translocates to the equatorial cortex and spindle midzone, eventually accumulating in the midbody during cytokinesis. Given the relevance to this project, I will outline here the known mechanisms governing Aurora B localisation to the centromere and to the spindle midzone and midbody.

TO THE CENTROMERE

The CPC first localises to chromatin in late S phase through INCENP binding to HP1 (Ainsztein et al., 1998). As mitosis progresses towards metaphase, the complex is released from arms by Aurora B's phosphorylation of H3S10, which opposes HP1 binding to the adjacent H3K9 methylation mark (Nozawa et al., 2010). The gradual focusing of the CPC to centromeres over prometaphase and metaphase is achieved via the recognition of two histone

marks: H2A phosphorylated at threonine 120 (H2AT120p) and H3 at threonine 3 (H3T3p). Respectively, these marks are laid down at centromeres in early mitosis by the kinetochore-associated kinase budding uninhibited by benzimidazoles (Bub1) (Kawashima et al., 2010) and haspin kinase (Dai et al., 2005). Phosphorylation of H3T3 creates a binding site recognised by the BIR domain of survivin (Kelly et al., 2010; Wang et al., 2010; Yamagishi et al., 2010). Structural studies have determined that the recognition of H3T3p by survivin is mediated by BIR domain recognition of an N-terminal tetrapeptide (Du et al., 2012; Jeyaprakash et al., 2011; Niedzialkowska et al., 2012). Bioinformatics searches using the tetrapeptide consensus identified shugoshin as a potential survivin interactor. The aforementioned H2AT120p mark is recognized by shugoshin that, in turn, interacts with borealin to recruit the CPC (Kawashima et al., 2007; Kawashima et al., 2010; Yamagishi et al., 2010). Localisation of the CPC to the centromere is particularly crucial for mitotic fidelity; the complex's role in this process is discussed later in this chapter.

TO THE MIDZONE AND MIDBODY

Upon anaphase onset, the CPC is first lost from the centromeres. CPC centromeric recruitment ceases as the histone H3 threonine 3 phosphorylation mark is removed from chromatin (Qian, 2011; Vagnarelli, 2011). Additionally, the CPC is actively removed from chromosomes as Aurora B is first ubiquitylated then removed by the CDC48 – an AAA+ ATPase (Ramadan, 2007). The subsequent targeting of the CPC to the central spindle during the transition into anaphase requires INCENP. Firstly, removal of a CDK1 mark on INCENP is required for re-location of the CPC to the spindle midzone (Hümmer and Mayer, 2009). The contribution of other CPC subunits to this localization is less clear. A survivin BIR-domain mutant that fails to

recruit the CPC to centromeres does not impede its relocalisation to the spindle midzone upon anaphase onset (Lens, 2006). Full Aurora B activity is required for CPC transfer to the spindle midzone at anaphase; cytokinesis remains defective in the presence of CPC hypomorphs, even those capable of rescuing chromosomal alignment and other defects associated with CPC loss of function (Xu et al., 2009).

ROLES OF THE CHROMOSOMAL PASSENGER COMPLEX

The CPC helps orchestrate numerous mitotic processes and the abrogation of its activity has pleiotropic consequences. In early mitosis, the CPC phosphorylates several chromatin substrates and is solely responsible for phoshorylating histone H3 on serine 10 and 28 (Giet and Glover, 2001; Sugiyama et al., 2002). The CPC also acts to increase recruitment of condensin I to mitotic chromosomes, correct erroneous kinetochore-microtubule attachments, activate the spindle assembly checkpoint and support spindle midzone formation. In cytokinesis, the CPC helps regulate contractile ring formation and abscission. The roles of the CPC in mitosis and cytokinesis are extensively reviewed in Carmena, 2012 and van der Waal et al., 2012. Due to their relevance to this project, I will elaborate here on Aurora B's role in chromosome bi-orientation and cytokinesis. It should also be noted that, further to the above mitotic and cytokinetic functions, a recent report suggests that Aurora B activity during interphase is required for proper mitotic progression (Hayashi-Takanaka et al., 2009).

Although not considered in great detail here, it is important to note that spatiotemporal direction and restriction of CPC activity is reinforced through a series of synergistic and antagonistic interactions with other mitotic regulators. These functional interactions include kinase-crosstalk with polo kinase (Lens et al., 2010), and the removal of phosphorylation marks

from Aurora B substrates by protein phosphatases (Emanuele et al., 2008; Liu et al., 2010).

IN SISTER CHROMATID BI-ORIENTATION

In order to maintain mitotic fidelity and ensure that each daughter cell receives one – and only one – copy of each chromosome, sister chromatids must be oriented to opposite poles prior to their separation. Failure to achieve bi-orientation can lead to aneuploidy, which can have serious consequences and potentially lead to cancer initiation (Holland and Cleveland, 2009). Thus, prior to their separation at anaphase, the attachment of sister chromatids to microtubules emanating from opposite poles must first be achieved and recognized by the appropriate sensors. Converging fields of Aurora B and protein phosphatase 1 (PP1) activity, respectively, emanating from the inner centromere and outer kinetochore, create a phosphorylation gradient across the kinetochore in metaphase. The PP1 gamma isoform is recruited to the outer kinetochore by its interaction with the N-terminus of KNL1, this interaction is itself opposed by Aurora B phosphorylation of the PP1 docking motif on KNL1 (Liu et al., 2010). This gradient promotes chromosome bi-orientation by destabilizing kinetochore-microtubule attachments that do not generate tension (Liu et al., 2009; Liu et al., 2010; Tanaka et al., 2002). Sister chromatid bi-orientation results in tension, which draws kinetochores outwards from the gradient of Aurora B activity. The importance of tension and the spatial segregation of Aurora B substrates in error correction has been illustrated in a study KMN network components, which are part of the outer kinetochore (Welburn et al., 2010).

Further to destabilizing incorrectly attached kinetochores to promote bi-orientation, Aurora B phosphorylation of several kinetochore components generates transiently unattached kinetochores that, in turn, activate the

spindle assembly checkpoint (Ditchfield et al. 2003, Hauf et al. 2003, Khodjakov and Pines 2010). As the presence of even a single unattached kinetochore is sufficient to produce a 'wait anaphase signal' (Rieder et al., 1995), the generation of transiently unattached kinetochores prevents premature anaphase onset. Incorrect (syntelic and merotelic) attachments that can be maintained in the absence of Aurora B function, satisfy the spindle checkpoint, leading to chromosome segregation errors (Hauf et al., 2003). Recent evidence also hints to a more direct role for Aurora B in checkpoint activation via the recruitment of the mitotic checkpoint kinase Mps1 (Saurin et al., 2011).

Rather than block mitotic progression, Aurora B inactivation can allow cells to proceed to anaphase without first achieving chromosome bi-orientation, resulting in widespread chromosome misalignment. It is for this reason that some choose to refer to Aurora B inhibitors as mitotic drivers rather than anti-mitotics (Keen and Taylor, 2009). Interestingly, a survivin BIR domain mutant that cannot localise the CPC to centromeres, but which does not affect CPC localization from anaphase onwards, can rescue lethality in the absence of wild type survivin in chicken DT40 cells (Yue et al., 2008). This mutant grew with comparable kinetics to cells containing wild type survivin but presented a small yet significant decrease in mitotic arrest when treated with taxol. Although functionally important, the biological importance of centromeric CPC-mediated error correction in vertebrates is uncertain.

IN CYTOKINESIS

Upon anaphase onset, the equatorial cleavage furrow ingresses, eventually leading to the formation of a midbody that will undergo abscission, resulting in the separation of the two daughter cells. Aurora B relocalizes from the centromeres to the spindle midzone and to the

equatorial cortex upon anaphase onset, where it persists as the midbody forms. Cytokinesis failure is a striking hallmark of CPC loss of function (Terada et al., 1998) but its role in this process is less well understood than that in chromosome bi-orientation.

The CPC interacts with a number of central spindle components and its activity is required for localisation of centralspindlin to the midzone (Douglas et al., 2010). Centralspindlin, in turns, regulates the recruitment of Ect2 – the guanosine exchange factor for RhoA (Nishimura and Yonemura, 2006). RhoA is a GTPase involved in contractile ring formation and maturation (Drechsel et al., 1997). Aurora B and INCENP also interact with a number of cytoskeleton components in several organisms, such as actin in *Dictyostelium* (Chen et al., 2007) and the formin FHOD-1 in humans (Ozlü et al. , 2010).

In *S. cerevisiae*, the CPC participates in the no-cut pathway, which acts to prevent DNA damage by delaying the completion of abscission should DNA be present in the cytokinetic furrow (Norden et al., 2006). Although experimental evidence is more limited, the CPC has been implicated in a similar pathway – known as the abscission checkpoint – in humans (Steigemann et al., 2009). The CPC interacts directly with components of the ESCRT-III complex via borealin, and Aurora B phosphorylates the CHMP4C component of the complex (Capalbo et al., 2012; Carlton et al., 2012). Phosphorylation of ESCRT-III components by Aurora B is thought to prevent the completion abscission thus regulate the timing of cytokinesis.

Cytokinesis failure leading to eventual mitotic catastrophe in subsequent mitoses is proposed to be the principal sequence of events by which Aurora B inactivation leads to cell death (Kaestner et al., 2009).

AURORA B IN CANCER

Aurora B overexpression has been reported in various cancer types and is associated with poor prognosis (Pannone et al.; Smith et al., 2005). However, Aurora B overexpression fails to transform cells in *in vitro* transformation assays (in contrast, Aurora A expression does transform cells, see (Wang et al., 2006) – it is, therefore, not an oncogene *per se* (Kanda et al., 2005). Furthermore, protein levels of Aurora B and CPC components peak in G2/M. As increased proliferation – thus elevated mitotic rates – are a hallmark of many neoplasms, higher levels of Aurora B may merely be coincidental in cancerous tissues. This would mean that an increased mitotic index, rather than increased levels of Aurora B, would correlate with poor prognosis.

AURORA B INHIBITORS AS ANTI-CANCER THERAPEUTICS

Despite the debate surrounding Aurora B's role in tumourigenesis, treatment with Aurora B inhibitors preferentially kills actively proliferating cells irrespective of p53 status (Ditchfield et al., 2003). Furthermore, the ATP analogues AZD1152 and GSK1070916, which preferentially inhibit Aurora B, have been shown to inhibit tumour growth in xenograft models (Hardwicke et al., 2009; Wilkinson et al., 2007). Taken together, these results have promoted Aurora B as an anticancer target (Girdler et al., 2006).

Further to cytokinesis failure, Aurora B inactivation permits mitotic progression in the presence of taxol-stabilized microtubules. Therefore, Aurora B inhibitors might potentiate the effects of taxane treatment by promoting premature mitotic exit and catastrophe, rather than mitotic arrest (Yang et al., 2009).

Several pan-Aurora and Aurora B-preferential inhibitors have undergone pre-clinical development and, in some instances, progressed to phase I and II clinical trials. Furthermore, a number of these compounds are now commonly used research tools – Aurora B-specific inhibitors are summarized in Table 1.3.

The current panel of Aurora B inhibitors consists solely of ATP-analogues. This raises two potential issues: lack of specificity and susceptibility to resistant mutants. Firstly, although these compounds show a marked affinity for Aurora B, these are promiscuous to a degree – inhibiting Aurora B paralogues at higher concentrations and, in some instances, unrelated kinases (see Bain et al., 2007) and the remark column in Table 1.3). Secondly, the common mechanism of action of Aurora B inhibitors is anticipated to limit its clinical targeting. A study carried out by the group of Stephen Taylor sought to identify Aurora B mutants resistant to ZM447439 (Girdler et al., 2008). These mutants readily arose in a hypermutable cancer cell line and, critically, proved to be resistant to all other Aurora B inhibitors tested. This result predicts the occurrence of pan-resistance to Aurora B inhibitors in the clinic and has been independently replicated in leukemia cells (Failes et al., 2012).

As alluded to whilst discussing antimitotics more generally, the performance of Aurora inhibitors in clinical trials has limited. In some instances, promising patient outcomes were observed but these are thought to have occurred due to the fortuitous binding of compounds to imatinib-resistant Bcr-Abl kinase (Fei et al., 2010; Modugno et al., 2007). The lack of biomarker(s) capable of identifying potential responder populations also

impedes the chances of success of Aurora B or pan Aurora inhibitors in the clinic.

Compound	Company	Phase	Indication	Remark
AZD1152	AstraZeneca	II	AML, advanced solid tumours	Aurora C inhibition
GSK1070916A	GSK	I	Advanced solid tumours	Aurora C inhibition
ZM447439	AstraZeneca	RT	N/A	Aurora A inhibition
Hesperadin	Boehringer Ingelheim	RT	N/A	LCK, CHK1, MAPKAP-K1, MKK1, PHK and AMPK inhibition

Table 1.3 – Currently available Aurora B preferential inhibitors, their clinical status (RT: research tool), indication (N/A: non applicable due to research use only), and remarks indicating any known off target effects (after Lens et al., 2010 and Carmena and Earnshaw 2003)

1.4 AIMS OF THE PROJECT

Evidence suggests that targeting Aurora B interactors might prove to be an effective alternative strategy for abrogating CPC function. A downregulator of survivin expression – YM155 – showed promise in inhibiting tumor growth in xenograft models but had very limited success in phase II clinical trials (Lewis et al., 2011; Nakahara et al., 2007). The potential of targeting protein-protein interactions within the CPC has been acknowledged but thus far dismissed due to perceived challenges in developing drugs capable of binding such surfaces (Keen and Taylor, 2004).

In light of the inability of certain INbox mutants to fully rescue Aurora B activity in the absence of endogenous INCENP (Xu et al., 2009), we propose the following hypothesis: competitive inhibition of the interaction between Aurora B and the INCENP INbox ought to result in loss of Aurora B activity and localization. Thus, this may constitute a novel strategy for the inhibition of CPC function. Because these would be distinct from the current crop of ATP-competitive inhibitors, PPI inhibitors of CPC function could be valuable for both research and clinical use. In theory, these would not be susceptible to Aurora B mutations conferring resistance to multiple ATP analogue inhibitors (Failes et al., 2012; Girdler et al., 2008).

Furthermore, INbox-mediated protein-protein interactions represent an attractive target as their disruption might curtail the activation not only of Aurora B, but also Aurora C – the latter is overexpressed in some cancers, which is associated with poor prognosis (Takahashi et al., 2000). Furthermore, ectopic Aurora C expression has the potential to compensate for loss of Aurora B function. Aurora C has been reported to rescue both Aurora B knockdown, and overexpression of catalytically dead Aurora B (Sasai et al., 2004; Yan et al., 2005).

The first aim of this project was to ascertain the effects of competitive inhibition of Aurora B-INbox binding. In contrast to previously published results, we wished to investigate conditions in which all endogenous CPC components are present. Should these results support the targeting of this interaction, a secondary aim of the project was to test whether this interaction can be modulated by drug-like compounds. That is, whether small compounds with drug-like properties can elicit the desired effect. For this, a library of small INbox mimetic peptides was designed, built and screened.

Current methods used to find short peptides and other small molecules disrupting PPIs rely on exogenous proteins, in either *in vitro* or in heterologous assay environments. In that setting, off-target effects are difficult to assess and the readout cannot include a measure of activity on endogenous substrates. SICLOPPS methodology was deployed to overcome the dual hurdles of peptide delivery and stability which currently hamper the assessment and optimization of short peptide mimics as antagonists of intracellular PPIs in cell-based assays.

As the intention was to assess peptide activity against endogenous proteins, new screening and library design approaches had to be considered. In contrast to their combinatorial counterparts, mimetic libraries would consist of a relatively small number of discrete members. Although less chemically diverse, discrete mimetic libraries are more amenable to cell-based screens using human cell lines and allow for complex screen readouts.

Assessing the potential of genetically encoded CP mimics to disrupt protein-protein interactions *in vivo* was a tertiary goal of this project. Given the wealth of structural and molecular biological data concerning the CPC upon which one can draw upon to intuitively design mimetic peptide inhibitors, Aurora B was deemed a strong candidate for trialing this approach.

CHAPTER 2: MATERIALS AND METHODS

2.1 TISSUE CULTURE TECHNIQUES

CELL CULTURE

HeLa Kyoto cells, a widely used adherent isolate of the HeLa cell line, were grown in Dulbecco's modified Eagle medium (DMEM; GIBCO, Life Technologies) supplemented with 1% fetal bovine serum (FBS, SIGMA) and 1X penicillin/streptomycin (GIBCO, Life Technologies). Cells were passaged as required by first washing in 1X phosphate buffered saline (PBS; GIBCO, Life Technologies), trypsinizing by the addition of TrypLE™ Express (GIBCO, Life Technologies) followed by a short incubation at 37^{o}C, and diluting in fresh warm medium. Typically, cells would be passaged at a 1/10 dilution every two days or upon reaching approximately 70% confluence.

For storage, cells were harvested from near confluent flasks by first washing in 1XPBS, trypsinizing and quenching with medium to a final volume of 10ml. Cells were transferred to a 15ml falcon tube, spun down gently (5 minutes at 1000 RCF), resuspended in 1.5ml freezing medium (10% dimethyl sulfoxide (DMSO), 90% FBS) and transferred to a cryogenic vial. Vials were stored at -80^{o}C for initial freezing and maintained at that temperature for short-term storage (1-3 months); for long term storage, vials were transferred to liquid nitrogen roughly one week after initial freezing.

TRANSFECTION-GRADE DNA PREPARATION

Unless otherwise stated, DNA samples used for transfection were prepared from 50ml overnight cultures using a (QIAGEN) as directed by the

manufacturer with the following modification: following initial filtration, 1ml endotoxin removal (ER) buffer (QIAGEN, available from the EndoFree kit range) was added to the cell lysates and incubated on ice for 30 minutes. After this step, cell lysates were applied to the QIAGEN-tip 100 and the remainder of the protocol followed as directed. DNA pellets were resuspended in 50µl ddH_2O at 4^oC overnight, DNA concentration was measured using a NanoDrop instrument (Thermo Scientific) prior to storage at -20^oC.

TRANSIENT TRANSFECTION

Cells were transfected using either FuGene 6 (Roche) or XtremeGENE 9 (Roche) according to the manufacturer's guidelines. It should be noted that Xtreme gene 9 is the replacement for FuGene 6, which was discontinued partway through this project; the formulation of these products is identical. Briefly, for HeLa cells in a 6 well plate: one day prior to transfection, cells were seeded to a density of 2.25x10^5 cells per well; cells were transfected using 2µg of plasmid DNA and 6µl FuGENE 6/XtremeGENE 9 diluted in 200µl Opti-MEM® (Life Technologies) following the manufacturer's instructions; 24 hour post-transfection, growth medium was changed and selection drugs added if applicable.

SUBCLONING

A HeLa Kyoto cell line already available in the lab, or a subclone thereof (HeLa Kyoto clone D, see section 3.3), was used throughout this study. The cell line was subcloned by seeding trypsinized cells to a density of <1 cell per well into a 96 well plate. That is, cells were diluted in supplemented medium to a density of 0.5 cells per 100µl, and 100µl of the mixture added to each well. The 96 well plate was incubated under usual

tissue culture conditions and wells regularly monitored for colony formation. After a week, cells from a number of wells containing a single colony were re-plated and expanded, then assayed for background abnormal nuclei indexes and transfection efficiency.

2.2 MOLECULAR BIOLOGY METHODS

SDS-PAGE

Prior to western blotting, protein samples were resolved by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Samples were diluted in sample buffer to 1X prior to loading. For HeLa cells, cells were lysed in 30µl 1X reducing SDS-PAGE sample buffer per $1x10^6$ cells, boiled for 5 minutes then sonicated for 15 minutes (settings: 30 seconds on/30 seconds off; high intensity) using a Bioruptor instrument (Diagenode). Lysates were cleared by centrifugation for 3 minutes at high speed prior to loading on a gel or storage at -20^{o}C. Lysates were resolved by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) in Laemmli buffer. For western blotting applications (see below), proteins were transferred onto Hybond ECL nitrocellulose membrane (GE Healthcare) by wet transfer at 4^{o}C overnight using a constant voltage of 30V.

PONCEAU S STAINING

Prior to western blotting, the extent and evenness of protein transfer were verified by Ponceau S staining. Following the disassembly of the transfer, membranes were first washed twice for 5 minutes in PBS-tween (1XPBS, 0.2% (v/v) Tween 20; hereby referred to as PBS-t) at room temperature with mild agitation. Membranes were then stained with Ponceau S solution (0.5% w/v Ponceau S, 1% acetic acid) for 10-15 minutes. Following staining, the Ponceau S solution was rinsed off repeatedly with

water until bands became visible over the background. All staining and washing steps indicated here were carried out at room temperature with mild agitation.

WESTERN BLOTTING

Following Ponceau S staining, membranes were washed twice in PBS-t and blocked in a solution of 5% (w/v) milk diluted PBS-t for 1 hour at room temperature. Blocked membranes were then washed 2x5 minutes in PBS-t then probed with the primary antibody diluted in 5% milk PBS-t for 1 hour at room temperature (see 2.3 for antibody dilutions used). Unbound primary antibody was washed off with 2x5 minutes PBS-t washes, then the desired secondary antibody diluted in 5% milk PBS-t was applied for 1 hour at room temperature. Prior to detection, the membrane was washed twice for 5 minutes in PBS-t, then a further two times for 10 minutes in PBS-t. When Li-Cor detection was used, the final wash was in detergent-free PBS to avoid the formation of bubbles on the membrane. Secondary antibody dilutions employed were of 1:5,000 for ECL detection, and 1:10,000 for Li-Cor detection.

2.3 MOLECULAR CLONING METHODS

AGAROSE GEL ELECTROPHORESIS

For DNA visualisation by agarose gel electrophoresis, the desired weight/volume of agarose (SIGMA) was melted in Tris-acetate-EDTA buffer (TAE). Gels were pre-stained by the addition of 0.5µg ethidium bromide (SIGMA) per mL of gel prior to casting. DNA samples were mixed with the appropriate volume of 6X gel loading dye (NEB) prior to loading and gels run at 100V until the dye front has reached the edge of the gel.

To resolve DNA fragments <1kb, 1% (w/v) agarose gels were typically employed and 0.8% (w/v) gels were used for larger fragments. In order to determine the size of DNA fragments, 0.5µg of either 1kb or 100bp DNA ladder (both NEB) were loaded alongside the samples to be analysed.

PCR AMPLIFICATION

Inserts that could not be directly inserted by simple restriction digest were amplified by polymerase chain reaction (PCR) with primers containing restriction site linkers using the Expand High Fidelity PCR system (Roche) as directed by the manufacturer. Annealing temperatures and the duration of the extension step were adjusted to suit the primer and PCR product size, respectively.

ENDONUCLEASE RESTRICTION DIGEST

All restriction enzymes used in this study were from NEB and used in the reaction buffers recommended and provided by the manufacturer, the reaction mixture was supplemented with 10mg/ml bovine serum albumin when recommended. The desired amount of DNA was digested using 10 units of enzyme per µg of DNA for 2 hours at 37^{o}C. When sequential digests were required, digested fragments were purified after the first digestion using the QIAquick PCR Purification Kit (QIAGEN) to remove the nuclease and buffer, then eluted in 30µl ddH_2O. The eluted DNA was then supplemented with the appropriate buffer and enzyme for the second digest.

BLUNT-ENDING DNA FRAGMENTS

When complimentary sticky ends could not be generated, both DNA fragments were first subjected to a single digest using a restriction endonuclease cutting in the desired location, then blunt ended. For 2-4µg of plasmid DNA, digested ends were blunted by the addition of 3 units T4 DNA

polymerase (NEB) supplemented with 10mg/ml bovine serum albumin and 400µM dNTPs, incubated at 12^{o}C for 15 minutes prior to heat inactivation at 75^{o}C for 10 minutes. Buffers and enzymes were replaced prior to carrying out the second, sticky-end generating, digest by using the QIAquick PCR Purification Kit (QIAGEN) and eluting the DNA with 30µl ddH_2O. The eluted DNA was supplemented with the necessary buffer and enzyme for the second digest.

PRIMER ANNEALING & LINKER LIGATION

In order to generate a myc-tagged Aurora B construct, a linker containing an ATG initiation codon followed a 3xmyc epitope tag plus upstream and downstream restriction sites was designed and inserted into the pIRES puro2 vector. Complimentary oligos containing an AflII site distal to the epitope tag were first annealed by diluting these to 100nM in ddH_2O and 10µl of each combined in a 1.5ml eppendorf to make 20µl of a 50nM solution. The annealing step was carried out in a large beaker of near-boiling water on a hot plate: the tube containing the oligos was added as the water neared boiling point and the beaker kept on the heat for 5 minutes prior to being removed from the head and allowed to slowly cool to ambient temperature. A foam float was used to keep the tube aloft in the water.

Although initially designed with ClaI and NheII compatible overhangs at the 3’ and 5’ end, respectively, the linker was eventually inserted into the destination vector using a 3’ blunt end and a 5’ AlfII-generated sticky end. These ends were generated by the blunting, then subsequent digest of 2µl the annealed linker mixture following the protocols outlined earlier. The modified linker was purified after each step using the QIAquick PCR Purification Kit (QIAGEN) and eluted in 30µl ddH_2O.

To phosphorylate the linker to allow ligation, in a PCR tube, 10 units of polynucleotide kinase (PNK; NEB) plus 2µl 10X T4 DNA ligase buffer (also NEB) were added to the purified annealed linker obtained following the previous step and made up to a final volume 50µl with ddH_2O. The phosphorylation mixture was incubated on a thermocycler for 30 minutes at 37^{o}C then inactivated at 65^{o}C for 20 minutes. The T4 DNA ligase buffer was used in lieu of the PNK buffer as it already contains the necessary concentration of unlabeled ATP. The phosphorylated annealed linker was ligated into a digested vector following the protocol described below and a 50:1 insert to vector ratio.

PHOSPHATASE TREATMENT

To limit vector re-ligation during cloning, digested vectors were treated with Antarctic phosphatase (NEB) following manufacturer's recommendations immediately following digestion. For 2-4µg of DNA, 5 units of Antarctic phosphatase were added to the digestion mixture and incubated at 37^{o}C for 30 minutes. The phosphatase was then inactivated prior to further handling by incubating the mixture for 5 minutes at 70^{o}C. Samples were then gel extracted following the protocol indicated below.

GEL EXTRACTION

DNA samples were mixed with Gel Loading Dye (6X) (NEB) to a final concentration of 1X loading dye and resolved by agarose gel electrophoresis on ethidium bromide pre-stained gels as described earlier in this section. DNA bands were visualized using a UV box and excised with a scalpel. Gel extractions were carried out using the QIAquick Gel Extraction Kit (QIAGEN) following the manufacturer's instructions and eluted in 30-50µl of ddH_2O, depending on the intended use. In instances where large quantities of DNA

were requited and the digestion protocol scaled up, the number of columns used during gel extraction was increased.

LIGATION

All ligations were carried out using the Quick Ligation™ Kit (NEB) as per the manufacturer's instructions. Briefly, a 1:3 ratio of vector and insert were mixed together and supplemented with 10µl 2X buffer and ddH_2O to a final volume of 19µl. 1µl of Quick Ligase was then added to the mixture and incubated at 25^{o}C for 5-15 minutes. To prevent prematurely exposing the competent cells to heat, the ligation mixture was chilled on ice for 2 minutes prior to transformation and amplification in *E. coli* – 10µl of the chilled ligation mixture were used to transform 100µl chemically competent Top10 cells.

TRANSFORMATION

For selection, storage and amplification, plasmids were transformed into *E. coli* as required. Aseptic technique and sterile reagents were used throughout. Unless otherwise specified, 0.5µl of a plasmid miniprep or equivalent quantities were used to transform 50µl chemically competent Top10 cells using the following protocol. Aliquots of Top10 cells were thawed on ice and the required amount of cells transferred to pre-chilled 1.5ml eppendorf tubes, after which the plasmid DNA was added to each tube and the mixture incubated on ice for 30 minutes. Cell were heatshocked at 42^{o}C for 30 seconds then cooled on ice for 2 minutes, after which the mixture was topped up with Luria-Bertani broth (LB broth) to 500µl. To allow expression of the selection marker prior to the application of selection, the transformed mixture was incubated at 37^{o}C with agitation for 1 hour. Cells were plated out onto LB plates containing the appropriate selection antibiotic. As a

negative control, plasmid DNA was substituted with ddH_2O – no colonies should be present in this sample following incubation on selective media.

SITE-DIRECTED MUTAGENESIS

Vectors were mutagenised by site-directed mutagenesis using oligos encoding the desired DNA base change(s). Using these oligos, the DNA was amplified from the vector using the high fidelity polymerase PfuUltra (Agilent Technologies) for 16 cycles. Following amplification, any remaining vector DNA present was removed by treating the DNA with 20 units DpnI (NEB) for 1 hour at 37^{o}C. The DpnI restriction endonuclease can cut methylated vector DNA but not the unmethylated amplification product generated using the mutagenic oligos. 1µl DpnI-digested DNA was used to transform 50µl chemically competent Top10 cells as indicated above. 50, 100 and 200µl of the 500µl recovery mixture were plated out on LB plates containing selection drugs, which were then incubated overnight at 37^{o}C. Selected clones were picked, miniprepped and analysed for the presence of the desired mutation and the integrity of the remainder of the open reading frame. Following site-directed mutagenesis, the mutagenized fragments were always subcloned back into the original vector by to ensure the integrity of the vector backbone.

SEQUENCING

DNA samples were sequenced as required using the Sanger (dideoxy) sequencing method. Briefly, 4µl of BigDye® Terminator reagent (Life Technologies) was added to 2µl of DNA prepared with a miniprep kit (or equivalent amount) and 2µl of 5µM primer, and made up to a volume of 10µl with ddH_2O. Samples were amplified for 25 cycles consisting of: 96^{o}C for 30 seconds, 50^{o}C for 15 seconds, then 60^{o}C for 70 seconds. All subsequent

sequencing reactions steps were carried out by the GenePool (now Edinburgh Genomics; Edinburgh, United Kingdom). Sequence analysis was carried out using Sequencher (Gene Codes Corporation).

2.4 MICROSCOPY METHODS

POLY-L-LYSINE COATED COVERSLIP PREPARATION

For microscopy analyses, coverslips were sometimes coated in poly-L-lysine to help retain mitotic cells. Glass coverslips were first acid washed in 70% ethanol, 1% hydrochloric acid for 10 minutes with gentle agitation. Coverslips were then rinsed in ddH_2O and dried on Whatman paper. Dried coverslips were then coated with poly-L-lysine by bathing in a 1:10 dilution of 0.1% (w/v) poly-L-lysine solution (SIGMA) in ddH_2O for 5 minutes. Coverslips were dried once again prior to sterilizing by washing in 70% ethanol and drying under UV light.

PHALLOIDIN STAINING

Coverslips were retrieved from experimentally treated wells and placed into wells containing warm 1X PBS. All subsequent washes and incubations were carried out at room temperature. Taking care not to disturb cells on the coverslips whilst aspirating liquids between washes, the coverslips were washed twice for 3 minutes with 1XPBS then fixed using 4% paraformaldehyde (PFA; Electron Microscopy Sciences) diluted in 1XPBS for 10 minutes. Following fixation, coverslips were washed twice for 3 minute with 1XPBS then samples were permeabilised by incubating the coverslips in 1XPBS, 0.2% Triton X-100 (Bio-Rad) for 5 minutes. At this juncture, the coverslips could be stored at $4^{o}C$ for several days prior to phalloidin staining.

Samples were stained with 0.5 units of either Alexa Fluor 488® phalloidin or rhodamine phalloidin (both Life Technologies) in 1XPBS, 1% (v/v) BSA for 20 minutes. Coverslips were then washed twice for 3 minutes in 1XPBS washes, briefly rinsed in ddH_2O, mounted using Vectashield with DAPI (Vector Labs) and sealed with nailpolish. Slides were stored in the dark at 4^oC pending analysis.

Representative images of experimental conditions were acquired using a Deltavision instrument (Applied Precision, GE Healthcare) using 40X and 100X oil objectives.

IMMUNOFLUORESCENCE

Unless otherwise indicated, wash and incubation steps were carried out at room temperature; blocking and antibody incubation steps were carried out on a parafilm support in a humid chamber using 100-50µl of solution. Experimentally treated coverslips were retrieved, fixed with paraformaldehyde and washed as per the phalloidin staining protocol above. Following fixation and washing, samples were permeabilised using 0.5% Triton X, 1XPBS for 5 minutes at room temperature then washed twice in 1XPBS for 5 minutes. Prior to antibody probing, coverslips were blocked in a solution of 3% BSA, 0.05% Triton X, 1XPBS for 1 hour then washed thrice in 1XPBS for 5 minutes. Primary and secondary antibodies were diluted in 0.5% Triton X, 1XPBS as indicated in table 2.3. Probing with primary and secondary antibodies were carried out in for 1 hour at 37^oC and 45 minutes at room temperature, respectively, and coverslips washed thrice in 1XPBS for 5 minutes after each step. Coverslips were then rinsed in ddH_2O and mounted on slides using VECTASHIELD® hardset mounting medium with DAPI (Vector Laboratories) and sealed with nail polish.

2.5 CPC FUNCTION ASSAY PROTOCOL

The protocol below was always performed using HeLa Kyoto clone D (described in Chapter 3) and was employed to obtain the results presented in Chapters 3 and 4, and the second pass screen in Chapter 5 (section 5.5).

Two days prior to transfection, cells were seeded to a density of $5x10^4$ in 6 well plates containing an uncoated 16mm diameter coverslip. Wells were transfected with 2µg of plasmid using 6µl FuGENE 6 (Roche) following the manufacturer's instructions. Transfected cells were selected by replacing the growth medium with fresh medium supplemented with 6µg/ml puromycin and incubated for a further 24 hours prior to collection. Each series of transfections contained two control wells transfected with empty pIRES puro2 vector to establish the background frequency of abnormal nuclei, and wells transfected with the $AURKB^{D218N}$ expression construct as a positive control.

Coverslips were harvested, washed, fixed and stained with phalloidin as indicated earlier. Coverslips were imaged on an Axioplan2 microscope using a Plan NEOFLUAR 40X/1,3 Oil objective (both Zeiss). ≥1000 cells were scored for the frequency of normal and abnormal nuclei. Abnormal nuclei indexes (ANI) were determined as the sum of multinucleate and micronucleated cells, and cells with chromatin bridges, divided by the total number of cells scored. The rate of increase in ANI relative to the control was calculated by dividing the ANI by the average frequency of abnormal nuclei detected in control coverslips transfected with an empty vector.

On occasion, the remaining cells in the well were harvested for western blot analysis. Wells were first washed with 1XPBS and the cells released by the addition of 0.5ml trypsin before being quenched with 0.5ml

of medium. Cells were then transferred to a 1.5ml eppendorf tube and pelleted in a benchtop centrifuge for 5 minutes at 800rpm. Pellets were washed by resuspending in 1ml warm 1XPBS, centrifuged once again and the supernatant removed prior to immediately freezing at -80^{o}C pending analysis. Pellets were lysed by the addition of 20µl 1XSDS sample buffer following the protocol indicated earlier; 20µl of lysate were loaded per gel lane on a 1mm 10 well 10% SDS-PAGE gel.

2.6 SICLOPPS LIBRARY PROTOCOLS

LIBRARY CONSTRUCTION

The INCENP INbox library was created using discrete HPLC-purified long oligos encoding the 97 library members purchased from Sigma-Aldrich (St Louis, Missouri, USA) – a table containing the library primer sequences can be found in Appendix II. Library inserts were amplified using Expand high fidelity polymerase (Roche) from a NcoI digested template using the PCR program described by Abel-Santos and colleagues (Abel-Santos et al., 2003). PCR products were purified using the QIAquick PCR purification kit (QUIAGEN) and eluted in 30µl ddH_2O prior to digestion with 10 units each of BamHI HF and AflII (both NEB) for 1hr at 37^{o}C.

Inserts were ligated into a phosphatase-treated BamHI and AflII digested pIRES puro2 vector containing the test SICLOPPS construct and a modified MCS lacking the AflII site. Ligations were carried out using as described earlier using a 1:5 vector to insert ratio. 10µl of the ligation mixture were used to transform 50µl Top10 chemically competent cells and 2/5th of the recovery mixture plated out on LB Amp plates and incubated overnight at 37^{o}C. Colonies were selected and screened for the presence of the desired insert by sequencing using the T7 forward primer.

In preparation for library screening, transfection-grade minipreps of all library constructs were prepared using the transfection-grade Plasmid Mini Kit (QIAGEN) following the manufacturer's instructions. DNA pellets were resuspended in 20µl ddH_2O overnight at 4°C, then quantified using the Nanodrop instrument and diluted to a concentration of 500ng/µl in ddH_2O. DNA samples were stored at -20°C. Protein-splicing deficient $SICLOPPS^{T69AH72A}$ mutants were generated by subcloning library inserts into the BamHI and AflII sites of a pIRES puro2-based SICLOPPS vector bearing said mutation downstream of the AflII site.

Unless modifications are indicated here, individual cloning steps were performed according to the protocols outlined in section 2.3.

LIBRARY SCREENING PROTOCOL – FIRST PASS

The initial library screen was carried out using a different protocol from the CPC assay. The chronological and practical reasons for employing this simplified protocol are discussed in Chapter 5.

One day prior to transfection, HeLa Kyoto cells were seeded to a density of $5x10^4$ in 6 well plates containing poly-L-lysine-coated coverslips. Duplicate wells were transfected with 2µg of plasmid each using 6µl FuGENE 6 (Roche) following the manufacturer's instructions. Transfected cells were selected by replacing the growth medium with fresh medium supplemented with 4µg/ml puromycin 24 hour post-transfection. Cells were then incubated for a further 24 hours prior to collection. Each series of transfections contained duplicate control wells transfected with empty pIRES puro2 vector to establish background frequencies of abnormal nuclei.

Coverslips were harvested and stained with phalloidin then imaged on an Axioplan2 microscope using a Plan NEOFLUAR 40X/1,3 Oil objective (both

Zeiss). For analysis, 5 fields of cells were scored per coverslip. Abnormal nuclei indexes (ANI) were determined as the sum of multinucleate and micronucleated cells, and cells with chromatin bridges, divided by the total number of cells scored. The rate of increase in ANI relative to the control was calculated by dividing the ANI by the average frequency of abnormal nuclei detected in control coverslips transfected with an empty vector.

2.7 BIOINFORMATICS RESOURCES

Alignments were performed with ClustalW (Thompson et al., 1994) using default settings, and visualised using Jalview (Clamp et al., 2004). Crystal structure visualisation and figure output were done using Pymol (Schrödinger, 2010). Microscopy images were handled with the softWoRx suite (Applied Precision) and ImageJ (Abramoff et al., 2004).

2.8 PLASMIDS USED IN THIS STUDY

Details of the SICLOPPS test construct donated by the Tyers lab can be found in Appendix I. The vectors pcDNA3 (Invitrogen, Life Technologies) and pIRES puro2 (Clontech) vectors were already available in the lab. Details of all the INbox library vectors can be found in Appendix II. Further details of the modified pIRES puro2 vector can be found in section 5.3.

Name	Vector	Selection	Details
SICLOPPS_test	pJ201	K	Original SICLOPPS vector
SICLOPPS^{WT}	pCDNA3	A; N	Test construct
SICLOPPS^{AA}	pCDNA3	A; N	Test construct; PD
GFP-LaminA	pCDNA3	A;N	
SICLOPPS^{WT}	pIRES puro2*	A; P	Test construct
SICLOPPS^{AA}	pIRES puro2*	A; P	Test construct; PD
AURKB^{WT}	pIRES puro2	A; P	N-terminal 3myc tag
AURKB^{D218N}	pIRES puro2	A; P	N-terminal 3myc tag kinase dead
INCENP^{WT}	pBluescript KS	A	Source of INbox
SICLOPPS_INbox^{WT}	pIRES puro2*	A; P	
SICLOPPS^{AA}_INbox^{WT}	pIRES puro2*	A; P	PD
SICLOPPS_INbox^{W845G}	pIRES puro2*	A; P	
SICLOPPS^{AA}_ INbox^{W845G}	pIRES puro2*	A; P	PD
SICLOPPS_INbox^{F881A}	pIRES puro2*	A; P	
SICLOPPS^{AA}_ INbox^{F881A}	pIRES puro2*	A; P	PD
SICLOPPS_INbox^{dbl}	pIRES puro2*	A; P	
SICLOPPS^{AA}_ INbox^{dbl}	pIRES puro2*	A; P	PD
SICLOPPS_INbox^{AAA}	pIRES puro2*	A; P	
SICLOPPS^{AA}_ INbox^{AAA}	pIRES puro2*	A; P	PD

Table 2.1 – Plasmids used in this study (WT: wild type; AA: T69AH72A mutant; dbl: W845GF881A double mutant; AAA: TSS→AAA phosphosite mutant*: modified pIRES puro2 vector; A: ampicillin; P: puromycin; N: neomycin; PD: processing deficient).

2.9 PRIMERS USED IN THIS STUDY

All primers used in this study were obtained through Sigma-Aldrich. Further to the primers listed in the table below, the sequence of all the individual primers used in library construction can be found in Appendix II.

Name	Application	Feature	Sequence
3myc_linkerF	linker	3xmyc; AflII	CGATGGAGCAGAAACTCATCTCTGAAGAAGA TCTGGAACAAAAGTTGATTTCAGAAGAAGAT CTGGAACAGAAGCTCATCTCTGAGGAAGATC TTAAGG
3myc_linkerR	linker	3xmyc; AflII	CTAGCCTTAAGATCTTCCTCAGAGATGAGCTT CTGTTCCAGATCTTCTTCTGAAATCAACTTTTG TTCCAGATCTTCTTCAGAGATGAGTTTCTGCT CCAT
T7	sequencing		TAATACGACTCACTATAGGG
AURKB_F	Cloning	AflII	CCGTCTTAAGATGGCCCAGAAGGAGAACTCC
AURKB_R	Cloning	EcoRI	CAGACCAGCCGAAGTTAGCAATCTTAGCTC
INbox_F	Cloning	NheI	GCCGGCTAGCCAATGGGGCTATTGCCCATAA TLGATCTGAATAGCGACGACTCCACCG
INbox_R	Cloning	AflII	TCTTCAAGAAGAGCAAGCCCCGCTATCACAA GCGCACCAGCTCTGAGTACTGCTTAAGCGGC
SICLOPPS_AA_F	SDM	T69AH72A	GGTTCTGTGAATAGAGCTGCTAGTGACGCTC GTTTCCTTACAACGGAC
SICLOPPS_AA_R	SDM	T69AH72A	GTCCGTTGTAAGGAAACGAGCGTCACTAGCA GCTCTATTCACAGAACC
AURKB_D218N_F*	SDM	D218N	GAGCTAAGATTGCTAACTTCGGCTGGTCTG
AURKB_D218N_R*	SDM	D218N	GGCAGAATTCGAGTCAGGCGACAGATTGAAG GGC
pIRES_AflIIMCS_F	SDM	~~AflII~~	CCGGTTAACAGGCCTATAGCGCTAGCTAGGC CGC
pIRES_AflIIMCS_R	SDM	~~AflII~~	GCGGCCTAGCTAGCGCTATAGGCCTGTTAAC CGG
INbox_W845G_F	SDM	W845G	GAAGCCCATCCCCACCGGGGCCCGAGGCAC
INbox_W845G_R	SDM	W845G	GTGCCTCGGGCCCCGGTGGGGATGGGCTTC
INbox_F881A_F	SDM	F881A	GGACTTGGAGGATATCGCCAAGAAGAGCAA GC
INbox_F881A_R	SDM	F881A	GCTTGCTCTTCTTGGCGATATCCTCCAAGTCC
INbox_AAA_F	SDM	TSS→AAA	GCTATCACAAGCGCACCAGCTCTGAGTACTGC TTAAGTTTTGG
INbox_AAA_R	SDM	TSS→AAA	CCAAAACTTAAGCAGTACTCAGAGCTGGTGC GCTTGTGATAGC

Table 2.2 – Primers used in this study. Restriction sites introduced by the primer are indicated in the 'Features' column; strikethrough denotes the removal of a restriction site (F: forward primer; R: reverse primer; SDM: site directed mutagenesis; IVT: *in vitro* translation; *: after (Girdler et al., 2006).

2.10 ANTIBODIES USED IN THIS STUDY

Antibody	Manufacturer	Species	Application	Dilution
α-myc (9E10)	Covance	mouse	WB	1:1,000
α-tubulin (B512)	Invitrogen	mouse	WB	1:15,000
ECL™ α-mouse IgG	GE Healthcare	sheep	WB – 2^o	1:5,000
α-mouse IRDye 800CW	Li-Cor	donkey	WB – 2^o	1:10,000
α-Aurora B (ab2254)	Abcam	rabbit	IF	1:500
α-tubulin (ATN02)	Cytoskeleton, Inc.	sheep	IF	1:200
ACA	N/A	human	IF	1:200

Table 2.3 – Antibodies used in this study. (N/A: non applicable, denotes patient sera or non-commercial antibodies available to our laboratory; WB: western blot; IF: immunofluorescence; 2^o: secondary antibody).

2.11 OTHER REAGENTS

BUFFERS

Name	Composition
3X SDS-PAGE sample buffer	50mM Tris-HCl pH 6.8; 15% sucrose; 2mM EDTA; 3% SDS (for 1ml 1X buffer: dilute in ddH_2O and add 90µl β-mercaptoethanol)
SDS-PAGE running buffer	25mM Tris pH 8.8; 192 mM glycine; 0.1% SDS
Transfer buffer	25mM Tris HCl pH8.8; 192mM glycine; 20% methanol
PBS	0.01 phosphate buffer; 137mM NaCl; 2.7mM KCl; pH 7.4
TAE	40mM Tris; 20mM acetic acid; 1mM EDTA
LB	1% tryptone; 0.5% yeast extract; 10mM NaCl; pH 7.4

Table 2.4 – General buffers used in this study.

SELECTION DRUGS

Selection drugs were freshly added to growth media as required. To avoid loss of efficacy over successive freeze/thaw cycles, small aliquots of the drugs were made prior to storage at -20^{o}C. All stock solutions were filtered through a 0.22µm filter prior to aliquoting.

Drug	Diluent	Stock concentration	Working concentration
Puromycin	DMSO	1mg/ml	1-3µg/ml
Ampicillin	ddH_2O	100mg/ml	100µg/ml
Kanamycin	ddH_2O	50mg/ml	50µg/ml

Table 2.5 – Selection drugs used in this study.

CHAPTER 3: SICLOPPS FUNCTIONAL TEST & A CELL-BASED CPC FUNCTION ASSAY

3.1 INTRODUCTION

As outlined earlier, one of the main aims of this project was to assess the effects of blocking the Aurora B-INCENP interaction, and test its amenability to being blocked with small compounds. Another aim of this project was to carry out a feasibility study to evaluate the use of SICLOPPS-generated cyclic peptide (CP) libraries in human cells. For this, we planned to assay a library of small SICLOPPS-generated CP mimics of the INCENP INbox in HeLa cells. Prior instances of intein-mediated peptide circularization in human cells in the literature were limited to Kinsella et al.'s study using the artificially split *DnaB* intein, We therefore first needed to test the functionality of the naturally-split intein we obtained for our study. The differences between theses two inteins are elaborated upon in section 3.3.

To assess the effect of small CP mimics of INCENP, it was also necessary to elaborate a cell-based assay capable of measuring a reliable read-out of CPC activity. This cell-based assay was refined over several iterations; the version of the assay presented here (sections 3.3) is the ultimate version used throughout this study with the exception of the first pass INbox library screen (section 5.4). The reasons for this exception are outlined in section 5.1. Discussions of the overall merits and limitations of the assay itself can be found in section 3.5 whereas finer grained discussion of the differences between this assay and the one employed in the first pass screen are confined to section 5.6.

3.2 RATIONALE FOR USING HELA CELLS

HeLa cells are transformed and possess a highly abnormal karyotype and numerous genetic lesions. As is typical of many transformed cells, HeLa cells posses a highly abnormal karyotype and are genetically unstable (Landry et al., 2013). HeLa cells do, however, possess many advantages that, although not unique, make them ideally suited for a cell-based screen of CPC function. They are adherent, easily transfected and susceptible to Aurora B inhibition (Ditchfield et al., 2003). For these reasons, HeLa cells were deemed a suitable model system and the expression and processing of the SICLOPPS test construct were assayed in the cell line.

3.3 FUNCTIONAL IN-VIVO TEST OF SICLOPPS

Prior to proceeding with library construction and screening, it was necessary to verify that SICLOPPS constructs can both be expressed and undergo post-translational splicing in human cells. At the time, the only report of intein-mediate peptide circularization in human cells was that made by Kinsella and colleagues in 2002, a study which employed an artificially split *DnaB* intein derived from *Synechocystis* sp. ATCC in human B lymphocytes. In its natural context, this intein exists as a single domain capable of excising itself from a protein; it has been artificially split to permit *trans* splicing (Wu et al., 1998).

To generate CPs, we opted to use a *DnaE* split intein-derived construct which, as described in Figure 1.1.1, is naturally present as two separate domains. This intein is frequently used for *in vivo* CP production and was being piloted for use in yeast by the Tyers lab within our institute at the time. This split intein had been demonstrated to undergo post-translational splicing in *Saccharomyces cerevisiae* but not yet in mammalian cells. A 3xmyc-tagged

SICLOPPS test construct, which is *S. cerevisiae* codon optimized, was kindly provided by R. Almeida and C. Arndt. An overview of the SICLOPPS test construct sequence and features is provided in Appendix I. I subcloned this construct into the *BamHI* and *EcoRI* sites of the CMV-driven mammalian expression vector pcDNA3 (Invitrogen). Insertion of the construct and sequence integrity were verified by sequencing (result not shown).

In order to ascertain that post-translational processing occurred, a negative control was required. The SICLOPPS^{H69AT72A} double mutant lacks residues that stabilize and orient the junction between the circularized fragment and the N-terminal intein, blocking cyclic peptide formation (Figure 3.3.1, panel B). This mutant has previously been used as a negative control for intein post-translational splicing (Kritzer et al., 2009) – please note that the T69 and H72 numeration refers to residue positions within the original N-terminal portion of the *DnaE* split intein, not that in the SICLOPPS orientation. I created this mutant by a single round of site-directed mutagenesis. The obtention of the desired double mutant was confirmed by sequencing; the integrity of the sequence of the full open reading frame (ORF) was also verified prior to subcloning back into the original pcDNA3 vector backbone. This was done to ensure that any mutation that might have been introduced outside of the ORF during site directed mutagenesis is removed.

To test SICLOPPS expression and processing, I first transiently expressed both SICLOPPS^{WT} and SICLOPPS^{T69AH72A} test constructs in HeLa Kyoto cells without selection. Whole cell lysates collected 48 hours post-transfection were analysed by western blotting using an anti-myc antibody recognizing the epitope tag affixed to the SICLOPPS precursor. Both the precursor and the faster migrating linear product (defined in panel A of

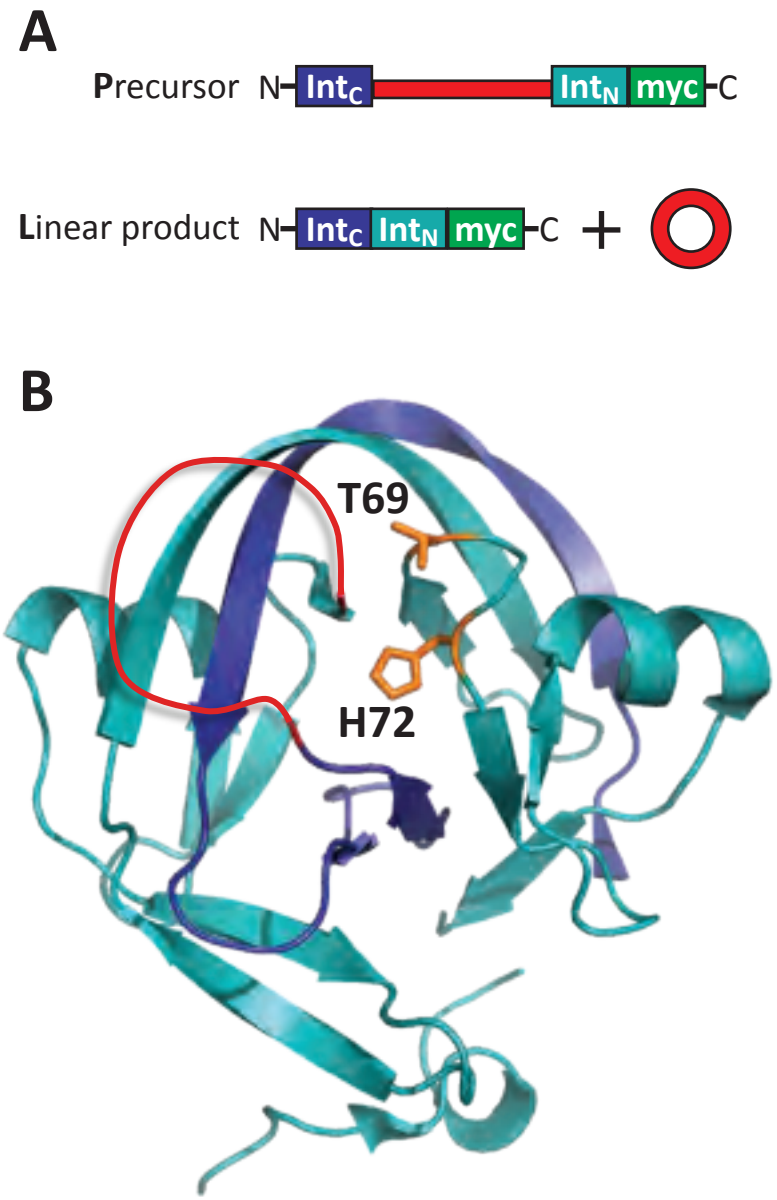


Figure 3.3.1 - (A) Post-translational processing of the myc-tagged SICLOPPS constructs used in this study yields a myc-tagged linear product (L) and an untagged cyclic peptide (red), a larger unprocessed myc-tagged precursor can also be detected. (B) Residues T69 and H72 (orange) of the N-terminal intein orient the C-terminal extein-intein junction and are necessary for intein processing. The colour scheme in this panel is the same one employed in panel A, the C-terminal myc tag is not show in panel B (both after Kritzer et al., 2009 and PDB ID: 1ZD7).

Figure 3.3.1) could be detected in SICLOPPS^{WT} sample, and the precursor only in the splicing-deficient control (Figure 3.3.2, panel A, lanes 4 and 3). These results are consistent with SICLOPPS processing occurring in HeLa cells. However, SICLOPPS protein levels appeared to be low as these required long exposures in order to be visualized using enhanced chemiluminescence (ECL) detection. The loading control was easily detected with short exposures, indicating that overall protein levels were sufficient, and that the transfer was efficient.

Wishing to ascertain whether the apparent low SICLOPPS protein levels were due to poor expression or low transfection efficiency, I sought to select transiently transfected cells. The pCDNA3 vector contains a neomycin resistance cassette conferring resistance to G418, an inhibitor of protein synthesis. This resistance marker is suited to stable, but not transient, transfection as selection typically takes 3 to 7 days of treatment in mammalian cells. I therefore subcloned the SICLOPPS test constructs into the *BamHI* and *BstXI* (blunted) sites of the multiple cloning site (MCS) of pIRES puro2 (Clontech). Like pCDN3, this vector contains a CMV promoter but also a puromycin resistance cassette, which is expressed from an *i*nternal *r*ibosome *e*ntry *s*ite (IRES). Puromycin inhibits protein synthesis and allows for the selection of transiently transfected cells within 24 hours of treatment. As the resistance cassette is present in the same transcript as the test construct, the translation levels from both ORFs should correlate, although these will not be identical.

The transient transfection experiment was repeated with the pIRES puro2 versions of the two SICLOPPS test constructs and 2µg/ml puromycin selection was applied for the last 24 hours. Bands corresponding to the SICLOPPS precursor and spliced products were once again present in the

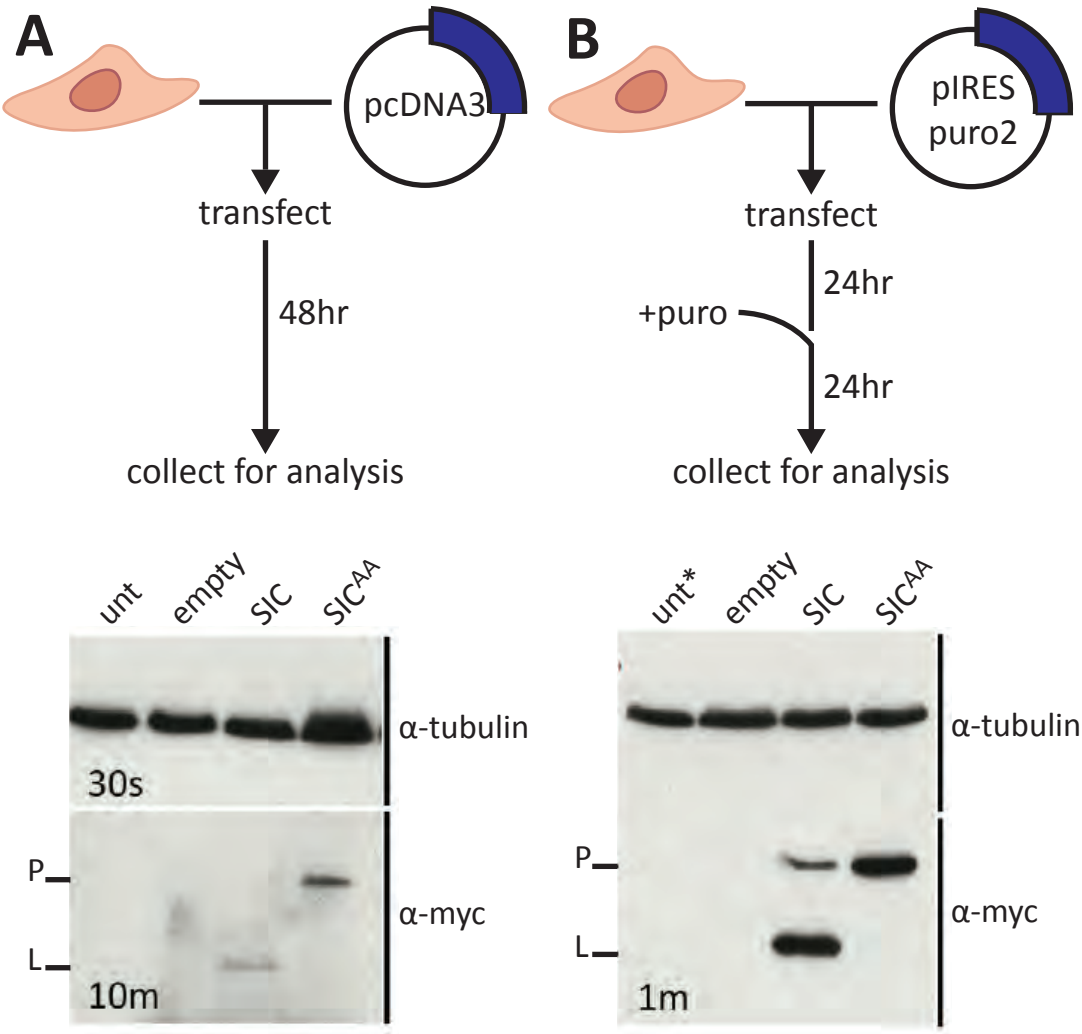


Figure 3.3.2 - SICLOPPS expression and processing tests in HeLa cells. Experimental schemes are outlined above their respective ECL immunoblots. Samples were either untransfected cells (unt), cells transfected with the empty vector, the SICLOPPS test construct (SIC) or the SICLOPPS^{T69AH72A} processing mutant (SIC^{AA}). (A) Experimental procedure and western blot analysis of cell lysates transiently transfected with pcDNA3-derived expression constructs. (B) As in panel A but performed using the pIRES puro2 vector along with puromycin selection prior to harvesting. (Exposure times are indicated on the bottom left-hand corner of each blot; loading density: 1.5^10^5 cells per lane; * denotes a sample not treated with puromycin; P and L labels correspond to the precursor and linear product as defined in Figure 3.3.1)

applicable samples 48 hours after transfection (Fig 3.3.2, panel B lanes 3 and 4, and lane 4, respectively). Moreover, these bands were readily detectable by ECL under the same conditions used in the previous experiment. Taken together, these results indicate that SICLOPPS constructs based on the *Ssp DnaE* split intein are readily expressed in HeLa cells, and undergo post-translational processing.

3.4 A CELL-BASED CPC FUNCTION ASSAY

CONSIDERATIONS

In developing a cell-based CPC function assay, the intention was not only to assess the effects of dissociating the Aurora B-INbox interaction (Chapter 4) but for it to form the basis of a genetically encoded library screen (Chapter 5). The library to be screened (section 5.2) was to consist of ~100 members. As the generation of such a high number of stable cell lines would be prohibitively time-intensive, this assay needed to be compatible with transient transfection.

For reasons elaborated in section 5.1, the initial phase of the library screen was carried out during piloting. In the interest of streamlining the number of factors to be optimized at that time, manual scoring by eye was favoured over automated scoring with or without automated image acquisition. This scoring method continued to be employed for the remainder of the study with additional handling steps to minimize scoring bias and permit statistical analyses (discussed further below and in section 5.6).

DESIGN & READ-OUT

When assaying genetic constructs by transient transfection, it is important to reliably identify transfected cells for scoring. Popular means of

achieving this are the marking of transfected cells (e.g. using a fluorescent marker) or the killing off of untransfected cells by antibiotic selection. As earlier results (section 3.3) indicated that puromycin selection worked well, this strategy was employed. pIRES puro2 was therefore used as a vector for all constructs presented in this section and throughout the remainder of this thesis.

Due to its multiple roles during cell division, loss of CPC function is known to have pleiotropic consequences affecting both chromosome segregation and cytokinesis (Adams et al., 2001). Furthermore, loss of CPC function leads to a loss of the Aurora B-specific histone H3 phosphorylation marks on serine 10 and 28 (Giet and Glover, 2001; Sugiyama et al., 2002). Of the consequences of CPC loss of function, cytokinesis is exquisitely sensitive – expression of even slightly hypomorphic INCENP INbox mutants fails to rescue cytokinesis defects in the absence of endogenous INCENP in chicken DT40 cells (Xu et al., 2010).

Population-wide cytokinesis defects lead to a rise in binucleate cells, which are easily recognizable by microscopy. Other nuclear morphological aberrations typical of CPC loss of function are the presence of micronuclei and chromatin bridges. Such features arise from lagging chromosomes, which are due to a failure to achieve chromosome bi-orientation prior to anaphase onset. Thus, the assay read out would encompass all these morphological nuclear aberrations (i.e. multinucleation, micronucleation, and chromatin bridges) as visualized by DAPI staining.

In order to better delineate cells and improve scoring accuracy, I also sought to highlight the cell perimeter. To achieve this, phalloidin staining was employed because it is an inexpensive reagent and the staining protocol is rapid. As 48 hours prove sufficient for SICLOPPS expression and processing,

and the selection of transiently transfected cells, this time point was chosen for the assay. The assay protocol is outlined in panel A of Figure 3.4.1.

SCORING METHODOLOGY & CRITERIA

Stained samples were scored for abnormal nuclei index (ANI; defined below) by eye using a 40X objective. To prevent operator bias during scoring, slides were first blinded and whole fields of view were scored until the total number of cells counted exceeded 1000. The rationale for the later step is that, should the investigator err towards first counting normal or abnormal cells/regions within a given field of view, the practice of stopping scoring as soon as 1000 cells are counted would result in a bias towards the investigator's scoring preference.

The ANI score is a measure of the frequency of cells within a sample presenting one or more of three nuclear morphological aberrations consistent with CPC loss of function: multinucleation, micronucleation, and chromatin bridges. Cells containing any of these three features, which are defined below and illustrated in Figure 3.4.1, as observed at 40X magnification were scored as having 'abnormal nuclei':

- *Multinucleation: cell con*tains ≥2 nuclei.
- *Micronucleation:* cell contains ≥1 micronuclei. Micronuclei are identifiable as round DAPI positive regions within the cell boundary; DAPI positive regions whose size exceeds ~3/4 that of the main nuclei were considered nuclei rather than micronuclei, cells presenting these were thus scored as multinucleated.

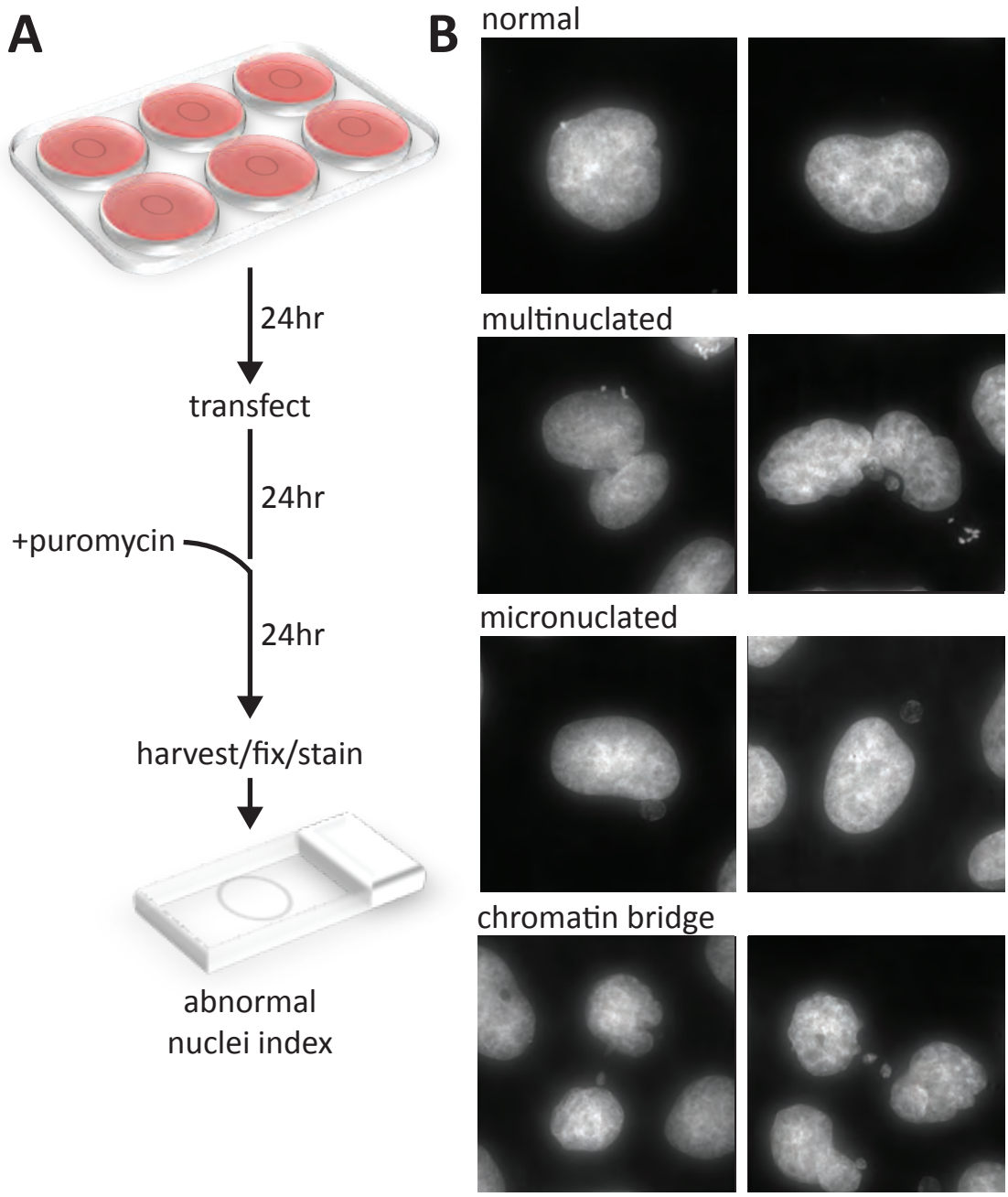


Figure 3.4.1 - (A) Outline of the CPC function assay protocol employed throughout all stages of this study with the exception of the first pass INbox library screen (chapter 5). (B) Representative DAPI stained nuclei for each of the scoring classes employed in the CPC function assay. For the purpose of this assay, multinucleated and micronucleated cells, as well as those containing chromatin bridges, are considered to posess abnormal nuclei. The abnormal nuclei index is the frequency of cells containing abnormal nuclei within the population.

- *Chromatin bridges:* cell contains a chromatin bridge. Chromatin bridges are identifiable as a single elongated DAPI-positive thread-like signal emanating from either the nucleus or a micronucleus towards to the junction between the two daughter cells; this signal is often mirrored in an adjacent cell.

Cells presenting abnormal features corresponding to more than one of the criteria outlined above were scored as abnormal only once, and their phenotype attributed to the upper-most category satisfied according to the following hierarchy: multinucleation > micronucleation > chromatin bridges. For example, a binucleate cell also containing a micronucleus would be scored as a multinucleate cell. Rare cells that could not be unambiguously categorized as either normal or abnormal were skipped during scoring.

RE-CLONING HELA KYOTO

The original HeLa Kyoto isolate used during screen piloting (see section 5.4) had a background ANI of 5.27 (SD=1.4%) under screen conditions (i.e. transiently transfected at low cell density with an empty vector and subjected to puromycin selection), and nearer to 7% under typical culture conditions. Prior to the second phase of the screen, we opted to re-clone the cell line in the hopes of obtaining a clone with a lower background ANI.

A fresh aliquot of low passage HeLa Kyoto cells was thawed and allowed to recover prior to isolating clonal colonies. Isolate colonies were first screened visually under a light microscope and those appearing to contain high ANI were discarded, as were those with noticeably fewer cells, which would suggest poor growth. Six clonal cell populations were then picked and expanded. To preserve cells at low passages, aliquots from each clonal population were made and stored in liquid nitrogen pending finer

characterization of these. Each clonal population was assayed for puromycin sensitivity, transfection efficiency, and their background ANI determined.

All the clonal populations assessed were efficiently killed by 4µg/ml puromycin in 24 hours (result not shown). In the initial screen 2µg/ml puromycin had been employed for selection but this was increased thereafter to further promote the detachment of dead (i.e. untransfected) cells from coverslips. Transfection efficiencies were determined by transient transfection using a GFP lamin A construct and visually scoring the frequency of cells with visible nuclear rim staining 24 hours post-transfection without any selection. A rough assessment of ANI without phalloidin staining was also made at this time by scoring >500 cells. Clones varied widely in their transfection efficiency but only mildly in their background ANIs (Figure 3.4.2, panel A).

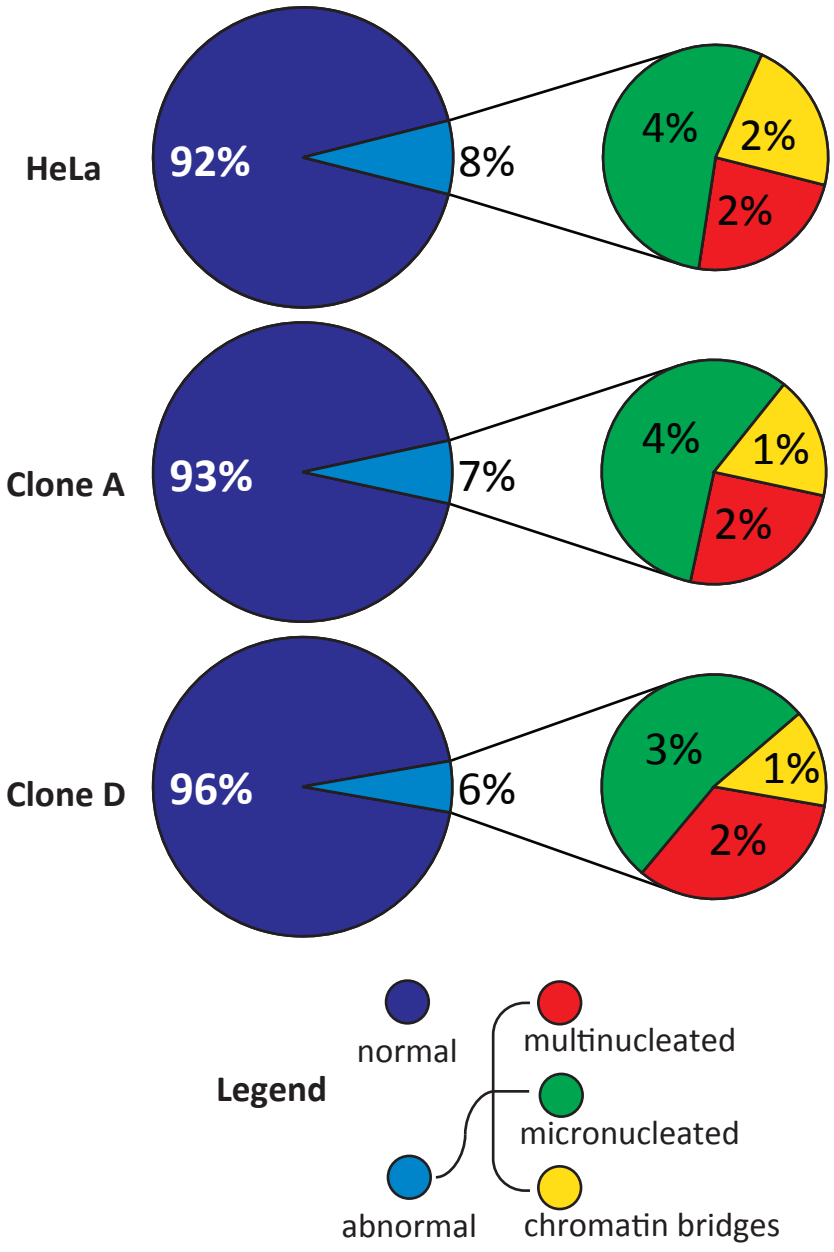
Clones A and D were selected for further investigation in light of their high transfection efficiency relative to the original cell line, and their low ANI. This consisted of determining their respective ANI based >1000 untransfected cells. Upon closer inspection, clone D had a marginally lower ANI than either clone A or the original HeLa Kyoto cell line (Figure 3.4.2, panel B). Due to its appreciably better transfection efficiency, Clone D was selected over the original cell line and employed in the cell-based assay. With respects to culturing this cell line, it was anecdotally observed during the first phase of library screening that background ANI tended to increase when cells were allowed to near confluency. Thus, this cell line was always passaged before reaching ~75% confluence.

A

Clone	Approximate ANI	Transfection efficiency
HeLa	7%	28%
A	7%	36%
B	8%	32%
C	8%	32%
D	7%	40%
E	9%	37%
F	6%	4%

Figure 3.4.2 - Recloning of the HeLa Kyoto cell line. (A) Rough abnormal nuclei index (ANI) and transfection efficiencies of the 6 isolated clones. Green boxes highlight clones selected for further analysis (see below) on the basis of their low ANI and high transfection efficiency. (B) Proportion of normal and abnormal nuclei present in HeLa Kyoto and the two best clones isolated quantified using a more stringent protocol (n>1000 cells). The clone designated 'Clone D' was used is in this study.

B



CONTROLS

In order to reflect the conditions of library screen, a genetically encoded positive control disrupting CPC function was desired. I thus opted to overexpress a dominant negative Aurora B kinase dead mutant. Two such active site mutants are reported in the literatures: K106R and D218N. The Aurora B D218N mutant was selected as it is reported to have the strongest dominant negative effect (Girdler et al., 2006). Inserts encoding wild type and kinase dead Aurora B (hereby referred to as $AURKB^{WT}$ and $AURKB^{D218N}$) were cloned into the pIRES puro2 vector. The Aurora B inserts were also N-terminally tagged with a 3xmyc epitope generated by linker ligation.

These constructs were assayed using the cell-based assay and their effect compared to that of an empty pIRES puro2 vector (Figure 3.4.3, panels A and B). In accordance with previously published results (Girdler et al., 2006; Honda et al., 2003), the effect of $AURKB^{WT}$ overexpression on ANI was indistinguishable from those of the empty vector control 48 hours post-transfection ($p>0.05$, n=3, Mann-Whitney U test). Overexpression of kinase-dead Aurora B had a very pronounced effect and elicited a highly significant 19.7 fold increase in ANI ($p<0.01$, n=3, Mann-Whitney U test). This result indicates that loss of CPC function is readily detectable within 48 hours using the cell-based assay devised.

3.4 SICLOPPS CONSTRUCTS DO NOT IMPAIR CPC FUNCTION

Further to testing the effect of dissociating the Aurora B-INbox interaction, we intended to use the cell-based assay to assay the effects of small CPs produced using the *Ssp DnaE*-based SICLOPPS backbone. It was therefore necessary to verify whether expression of this backbone itself interferes with the assay. For this purpose, the effects of the $SICLOPPS^{WT}$ and

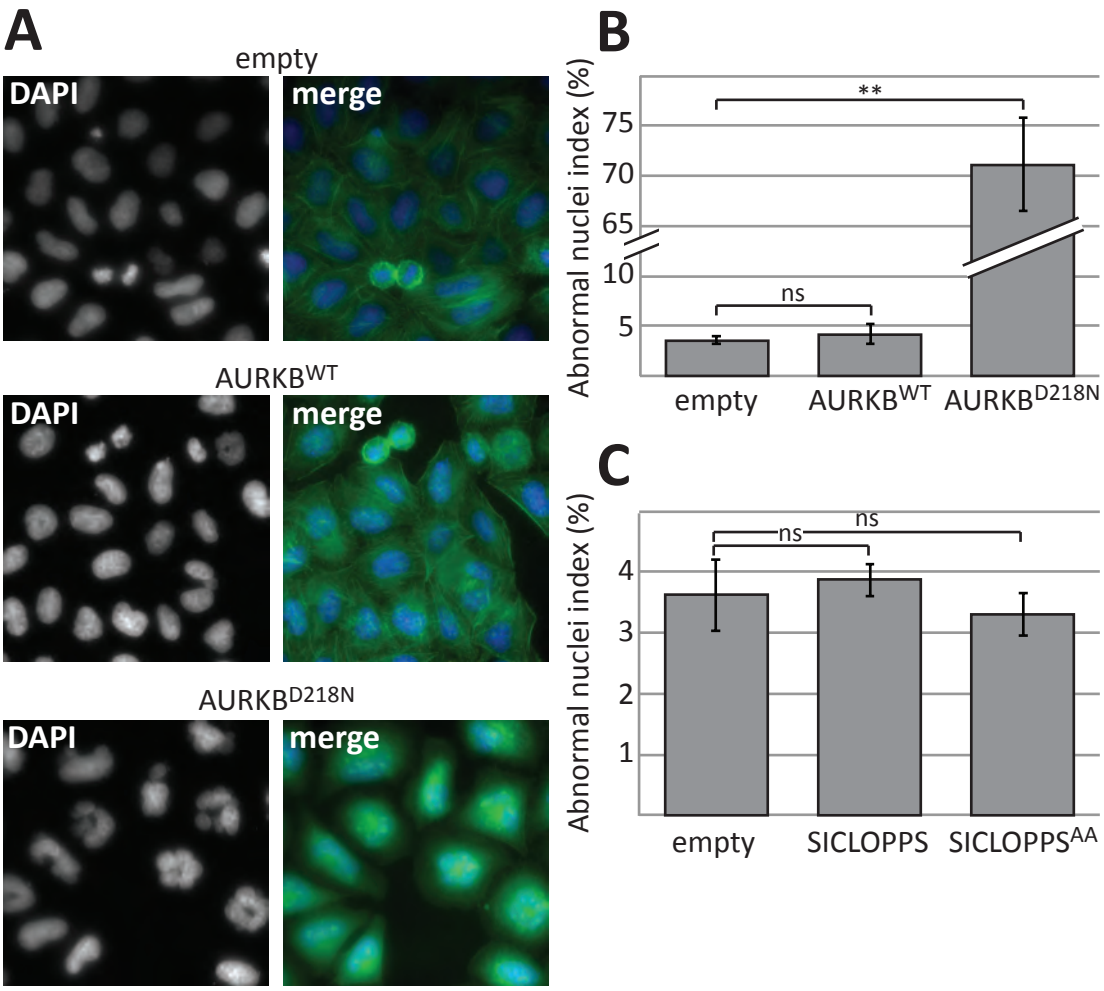


Figure 3.4.3 - CPC function assay controls. (A) Representative micrographs showing transiently transfected DAPI (blue) and phalloidin (green) stained HeLa Kyoto cells overexpressing wild type ($AURKB^{WT)}$ or kinase dead ($AURKB^{D218N)}$ Aurora B 48 hours post-transfection (40X magnification). (B) Quantification of the abnormal nuclei index (ANI) in experimental samples treated as in (A). (C) Quantification of the ANI in samples overexpressing the SICLOPPS test construct or a splicing-deficient mutant thereof ($SICLOPPS^{AA}$). In panels B and C: n=3, >1000 cells per replicate, error bars +/- SD, **: $p<0.01$, ns: non significant.

SICLOPPS^{T69AH72A} test constructs employed earlier (section 3.3) were assayed using the protocol devised above. The extein region of these constructs, designed by the Tyers laboratory, encodes the peptide CFGGSGGHPQFANA, which has no primary amino acid sequence homology to CPC components.

As illustrated in panel C of Figure 3.4.3, upon expression for 48 hours – the last 24 hours of which under puromycin selection – neither SICLOPPS^{WT} nor SICLOPPS^{T69AH72A} construct elicited a significant change in ANI (both $p>0.05$, Mann-Whitney U test). SICLOPPS libraries were thus deemed compatible with the cell-based assay.

3.5 DISCUSSION

SICLOPPS FUNCTIONAL TEST

SICLOPPS test constructs based on the *Ssp DnaE* split intein were readily expressed and processed in HeLa cells. To our knowledge, this is the first instance of the *Ssp DnaE*-derived SICLOPPS construct being shown to undergo post-translational splicing in human cells. Since these results were obtained, a study was published demonstrating that the naturally-split *DnaE* intein from *Synechococcus elongatus* is capable of *trans* splicing in human cells (Chen et al., 2010).

It should, however, be cautioned the CP product was not directly detected. Instead, I infer the production of the CP on the basis of the formation of the linear product of intein processing. Due to their small size and lack of epitope tag, it is difficult – if at all possible – to detect CPs themselves by immunoblotting. In fact, there are no published reports of western blot detection of SICLOPPS derived CPs, only larger circularlised proteins. The *in vivo* production of a 12 residue long CP generated in *S. cerevisiae* using the same intein employed in this study has been confirmed by mass spectrometry analysis of purified cell lysates by Kritzer and colleagues (Kritzer et al., 2009). Due to time constraints, we have opted not to reproduce this experiment in human cells.

RECLONING OF THE HELA CELL LINE

The HeLa cell line and its derivatives are transformed and genetically unstable. In hindsight, the attempt to reclone HeLa Kyoto was therefore unlikely to yield a clone with an appreciably lower ANI, or one that would remain stably so. Nonetheless, a clone was obtained that had an improved transfection efficiency compared to the starting HeLa Kyoto cell line. This

clone was used for most of this study and it was therefore necessary to explain its origins in this report.

CPC FUNCTION CELL-BASED ASSAY

The phenotypic assay needed to satisfy two key criteria for its use as a SICLOPPS-derived library screening platform: sensitivity to loss of CPC function, and insensitivity to SICLOPPS constructs themselves. Whilst the lack of ANI increase indicates uncompromised CPC function, the inverse (i.e. that an ANI increase *must* be due to compromised CPC function) may not necessarily be true. As such, one must exercise caution in attributing, be it directly or indirectly, an observed ANI increase to impaired CPC function solely on the basis of the cell-based assay described here. Doing so will require additional evidence.

Here, I will evaluate several features of the CPC function assay in comparison to alternatives that could have been employed. I will then expand on a number of foreseeable potential limitations of the assay.

Cytokinesis failure results not only in nuclear morphological aberrations but also in increased DNA content. An alternative screening strategy considered was to employ flow cytometry to detect increased DNA content. A particularly attractive possibility was that this would, if used in conjunction with a fluorescent detection of the epitope tag, permit an assessment of the relationship between transgene expression levels and DNA content. However, this strategy was not pursued due to the difficulties in distinguishing between mitotic cells and those having failed cytokinesis – both have 4N DNA content. Such an approach could, however, be useful in validating hits.

The use of puromycin selection to enrich for transfected cells is well suited given the manual scoring employed here (i.e. scoring by eye without image capture). Although suitable, under these scoring conditions, the use of a fluorescent tag (e.g. GFP) to identify transfected cells might give rise to a scoring bias favouring the most highly expressing (i.e. brightest) cells. As the puromycin resistance gene is expressed from an internal ribosome entry site (IRES), the expression of the marker should – although being weaker – correlate with that of the transgene. Thus, a minimum transgene expression level can be imposed through the concentration of puromycin applied during selection. However, there may be significant expression heterogeneity above that threshold.

An additional nuclear phenotype that could have been included in this assay, and might have been informative, is one encompassing cells presenting 'grape-like' clusters of small nuclei. This morphological feature is indicative of mitotic slippage in which the cell exists mitosis in the presence of scattered chromosomes. The reformation of the nuclear envelope around these scattered chromosomes results in the formation of multiple small nuclei. For the purpose of this assay, such cells were scored as being multinucleated. Although not quantified, this phenotype was frequently observed in cells overexpressing kinases-dead Aurora B.

Given that constructs are transiently transfected, vector dilution over successive cell cycles limits the time period over which these can be assayed. A 48 hour assay time point was selected to allow sufficient time for puromycin selection and for cells to undergo several cell cycles whilst remaining in the presence of assay construct. Furthermore, following cytokinesis failure negatively impacts on cell survival. Thus, longer assay periods might also risk resulting in the gradual depletion of the most severely

affected cells. The effects of the AURKB^{D218N} positive control were readily apparent 48 hours post-transfection. This result indicates that this period of exposure is sufficient for assaying the effects of genetically encoded constructs on CPC function.

CHAPTER 4: EXPRESSION OF SOLUBLE INBOX IMPAIRS CPC FUNCTION

4.1 INTRODUCTION

In order for the CPC to perform its many functions, Aurora B must not only be active but also correctly localized. Loss of either, or both, of these features has serious deleterious consequences. Loss of Aurora B activity through active site inhibition results in chromosome segregation errors, cytokinesis failure and other mitotic defects (Ditchfield et al., 2003).

Loss of localization can be achieved through different means. Treatment of cells with actinomycin D – an antibiotic known to inhibit transcription and bind DNA duplexes – causes mislocalisation of the active and intact CPC by an unknown mechanism (Becker et al., 2010). A number of INCENP and survivin mutations disrupting either the survivin BIR domain or the formation of the triple helix bundle supporting the localization module have been reported to block CPC targeting to centromeres but not to the spindle (Lens et al., 2006; Yue et al., 2008). In chicken DT40 cells, survivin^{D72AD73A} can rescue loss of endogenous survivin despite failing to localize the CPC to the centromere from prophase to metaphase. Such cells, however, present a mild but significant checkpoint defect. Depletion of CPC targeting components (i.e. INCENP, survivin and borealin) disturbs both Aurora B activation and localization throughout mitosis, leading to chromosome segregation and cytokinesis defects (Adams et al., 2001; Carvalho et al., 2003; Gassmann et al., 2004; Lens et al., 2003). Chemical inhibition of Aurora B causes similar defects (Hauf et al., 2003). It is important to note that depletion of INCENP, borealin or survivin tends to result in a

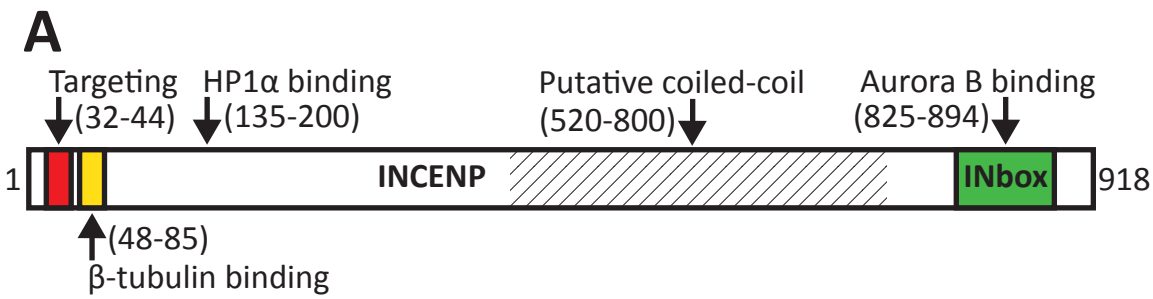
concomitant drop in the protein levels of the other CPC regulatory components (Klein et al., 2006).

I hypothesized that simply abrogating the interaction between INCENP and Aurora B would significantly impair CPC function, presumably through the loss of Aurora B activation and localisation. Aurora B localisation would be disturbed due to loss of linkage not only with the localization module of the CPC consisting of borealin, survivin and the INCENP C-terminus, but also with the INCENP flexible coiled domain required for spindle association (Figure 4.2.1, panel A). This might, in turn, quell activation by preventing Aurora B clustering.

To date, the experimental evidence most relevant to the targeting of this interaction is limited to the depletion or knockout of CPC components. Although highly informative, this experimental approach cannot fully recapitulate the effects of dissociating the INbox and Aurora B interaction alone, due to the loss of the endogenous protein. Thus, it was uncertain that the consequences of simply abrogating this interaction would reflect these experimental results. The effects of the loss of this interaction in the presence of all intact endogenous CPC components were of particular interest.

4.2 EXPRESSION OF SOLUBLE INBOX IMPAIRS CYTOKINESIS

To test the effects of disrupting the Aurora B-INbox PPI, I opted to express the INbox domain alone. In isolation, this fragment of INCENP should retain the ability to bind Aurora B whilst lacking that to associate with microtubules, survivin and borealin. I sought to overexpress the full INbox domain of human INCENP in HeLa cells using a *Ssp* dnaE-derived SICLOPPS scaffold for circularisation to confer stability (detailed in Chapter 3). An insert



D

Mutant	Residue(s)	Location	Mutant
F845G	F	845	G
W881A	W	881	A
dbl	F;W	845;881	G;A
AAA	TSS	892-894	AAA

Figure 4.2.1 - (A) Domain architecture of INCENP highlighting regions of interaction with other proteins. The targetting domain (red) refers to the region participating in the triple helix bundle with borealin and survivin. (B) Amino-acid sequence of human INCENP residues 821 to 900, which contain INbox (green); the region of the INbox stably associated with the Aurora B N-terminal lobe in the *Xenopus laevis* crystal structure is contained within the boxed region. INbox residues known to be critical for CPC function are highlighted in white. (C) Sequence of the 75 residue INbox peptide expressed using SICLOPPS. Accessory residues include the +1 cysteine required for intein processing (orange) and native extein residues (blue) included to promote processing. Residues highlighted in white were mutated in some experimental constructs as indicated in the table in panel D.

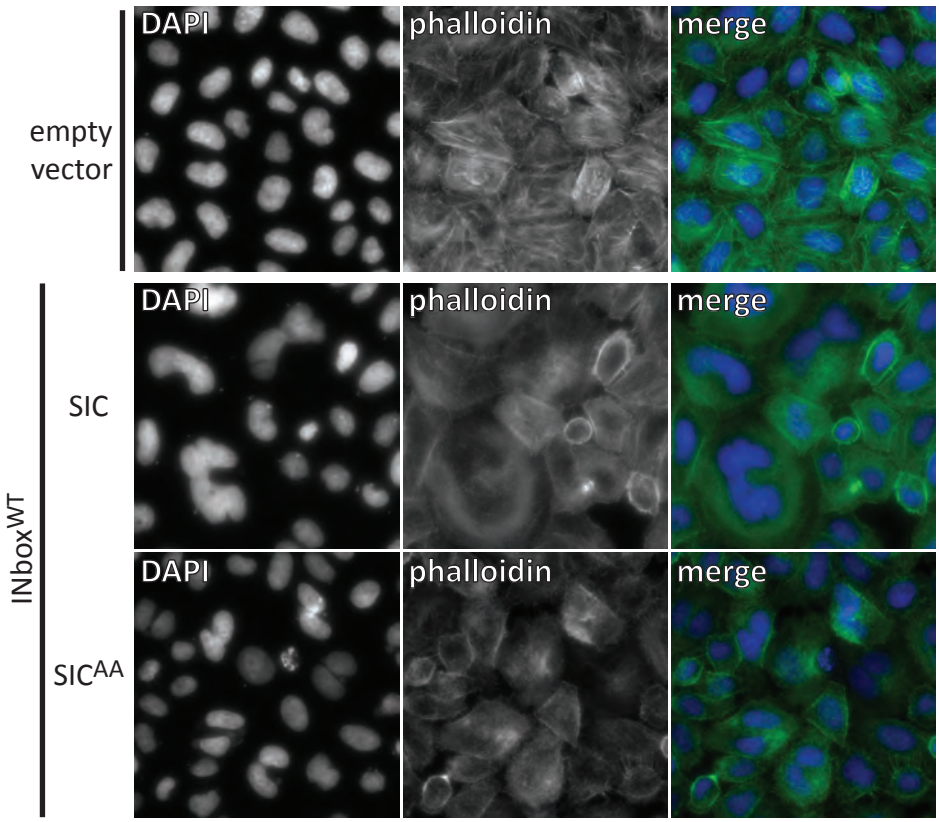


Figure 4.2.2 - Overexpression of the INCENP INbox causes an increase in cells containing abnormal nuclei. HeLa cells were transiently transfected with splicing-competent (SIC) and deficient (SIC^{AA}) SICLOPPS constructs containing the human INbox sequence in the extein position. Representative DAPI and phalloidin stained micrographs of samples collected 48 hours after transfection, the last 24 of which under puromycin selection, as imaged using a 40X objective.

spanning residues 821 to 900 of human INCENP – hereby refered to as INbox^{WT} – was inserted in the extein position of the pIRES puro2 SICLOPPS^{WT} and SICLOPPS^{T69AH72A} test constructs (Figure 4.2.1, panel B). In the hopes of enhancing processing, which can be sensitive to extein residues(Scott et al., 2001), accessory residues were added to either end of the INbox insert. Native extein residues found in the *Ssp DnaE* protein were used as these are known to support the processing of the intein. Further to the cysteine residue in position +1, which is required for processing, the insert was flanked with the +2 and +3 native C-terminal, and -1 and -2 N-terminal extein residues (Figure 4.2.1, panel C). Thus, processing of this construct is expected to yield a 75 amino acid residues long CP spanning the whole INbox.

This construct was assayed using the CPC function assay described in the previous chapter. Transient expression of this construct in HeLa cells had a significant dominant negative effect that could be readily observed by microscopy (Figure 4.2.2). This effect was quantified and the extent of SICLOPPS processing measured by quantitative western blot analysis. To permit side-by-side comparison, the results of these analyses with different INbox constructs are compiled in Figure 4.4.2 at the end of this chapter. 48 hours after transfection, a 5.1 fold increase in abnormal nuclei index (ANI) could be observed relative to the empty vector control (Figure 4.4.2, panel A; $p<0.05$; Mann-Whitney U test, n=3). Western blot analysis of cell lysate revealed that the extent of SICLOPPS processing was of approximately 30% (Figure 4.4.2, panel B). Western blotting analysis indicated that the extent of SICLOPPS processing was comparable across all INbox SICLOPPS constructs assayed as part of this study (Figure 4.4.1, panel B, and data not shown for INbox^{AAA}).

As post-translational processing of the *Ssp dnaE* intein was only partial, both the circularized INbox and a linear INbox flanked by the two intein halves (hereby referred to as the linear precursor) should be present. To discriminate between the effect of the cyclic INbox peptide and its linear precursor, a SICLOPPS^{T69AH72A} processing deficient version of the INbox construct was generated in the same vector and assayed. The lack of intein processing was confirmed by western blotting (Figure 4.4.2, panel B). Unexpectedly, the SICLOPPS^{T69AH72A} mutant with the INbox^{WT} insert produced a more pronounced ANI increase relative to its processing-competent counterpart. Transient expression of the splicing-deficient form of the INbox^{WT} caused a 7.9 fold increase in ANI (Figure4.4.2, panel A; $p<0.05$; Mann-Whitney U test, n=3). The relative decrease in ANI between the processing-deficient and competent form of the INbox construct is of about one third, which is roughly the same as the extent of intein processing in the SICLOPPS construct. These results suggest that it the linear precursor is predominantly responsible for the phenotype observed.

4.3 INBOX RESIDUES W845 AND F881 ARE NECESSARY FOR THE DOMINANT NEGATIVE EFFECT OF THE SOLUBLE INBOX

Previous work has highlighted several highly conserved features of the INbox that are absolutely necessary for its function. Residues W845 and F881 of the human INCENP are the most highly conserved INbox residues and make close contacts with the N-terminal lobe of Aurora B in both the *Xenopus* and human crystal structures of the Aurora B:INCENP complex (Elkins et al., 2012; Sessa et al., 2005). Mutants in which the residues homologous to W845 and F881 are substituted for others are reported not to bind nor activate, and not to activate Aurora B, respectively. In accordance with Xu et al., 2009, residue W845 was substituted for a glycine, and F881 for

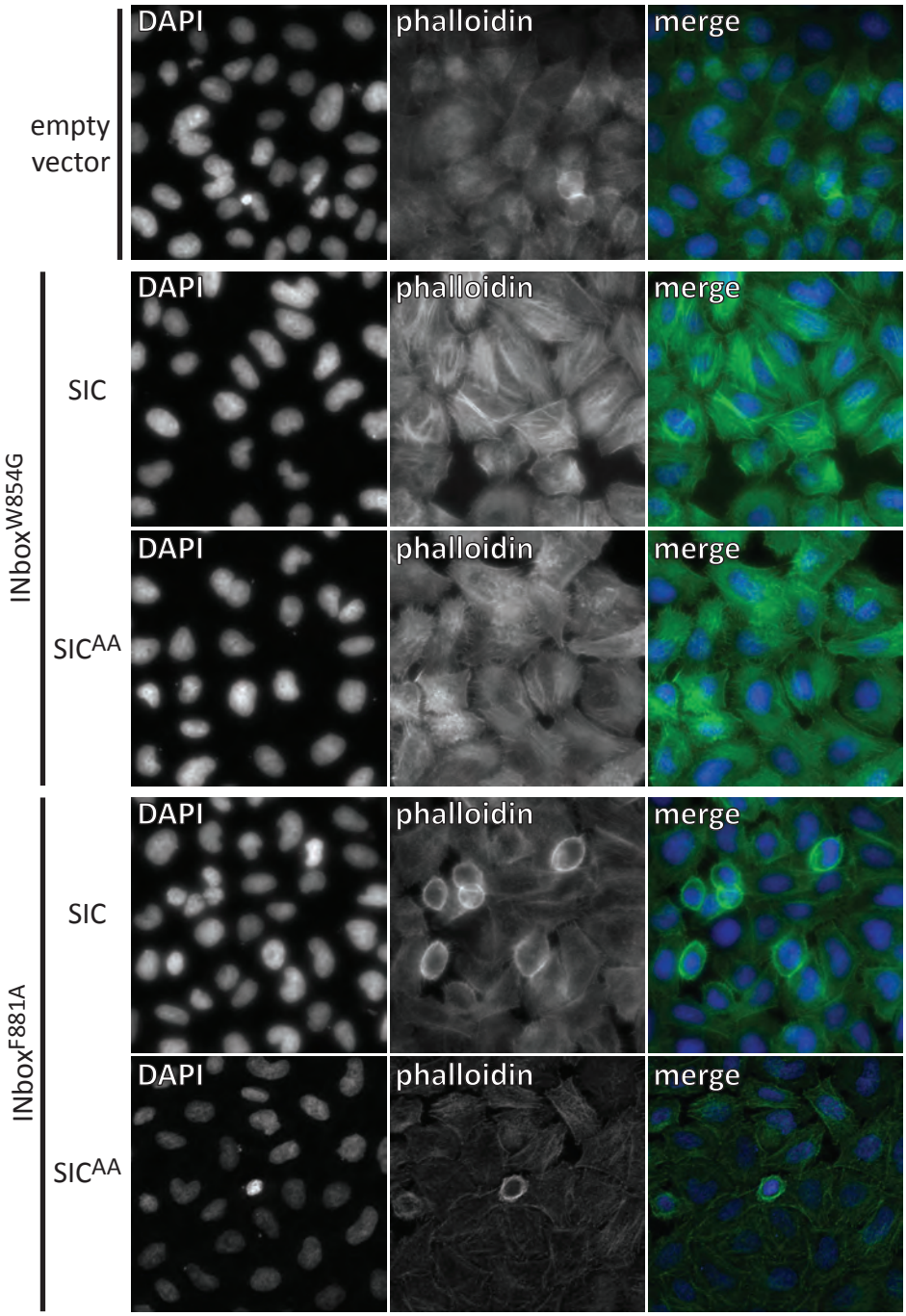


Figure 4.3.1 - Overexpression of INCENP INbox mutants lacking W845 or F881 does not impair CPC function. Representative DAPI and phalloidin stained micrographs of transiently transfected samples treated as in Figure 4.2.2.

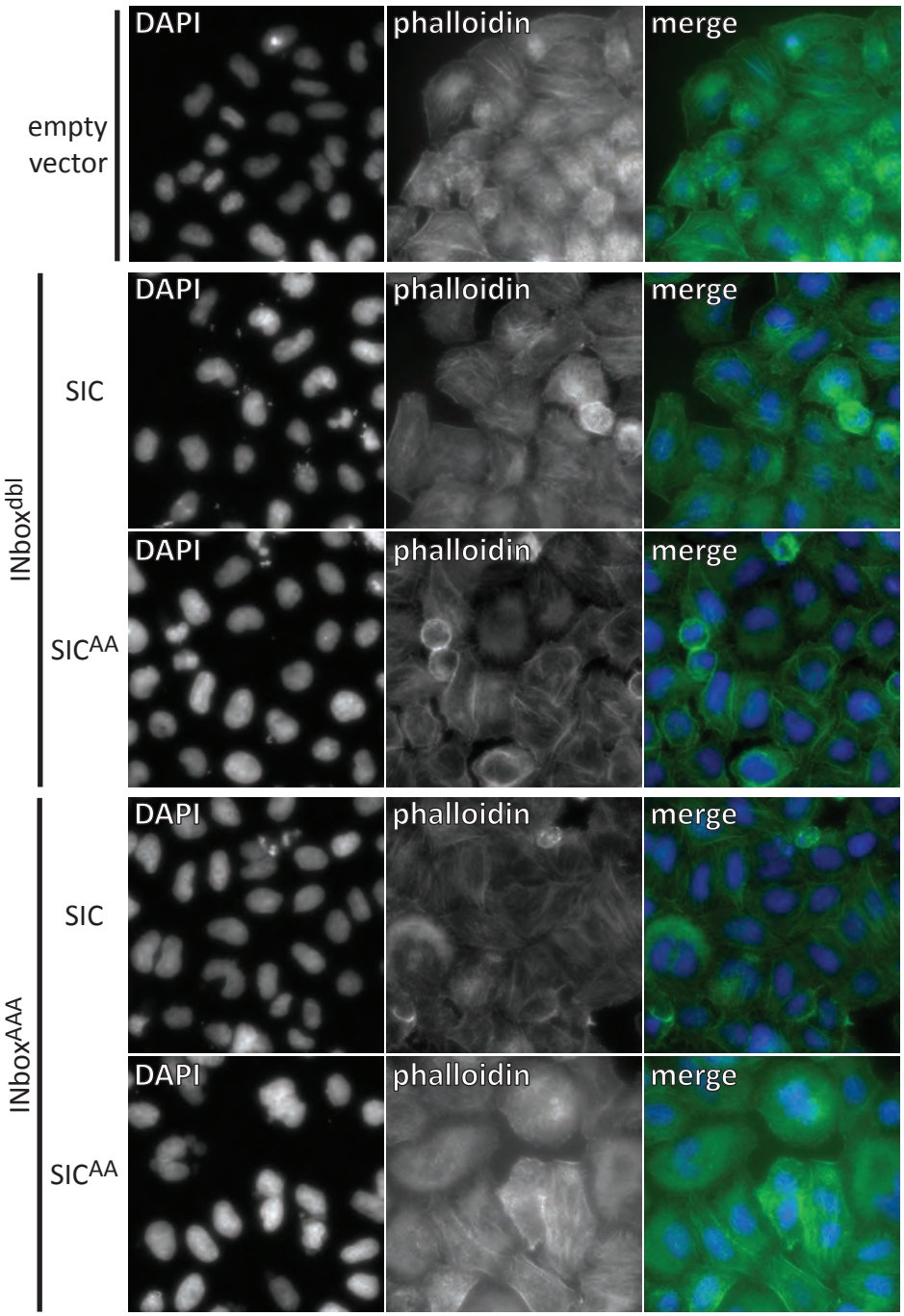


Figure 4.3.2 - Overexpression of a INCENP INbox TSS motif mutant, but not one lacking W845 and F881, impairs CPC function. Representative DAPI and phalloidin stained micrographs of transiently transfected samples treated as in Figure 4.2.2.

an alanine in the two versions of the SICLOPPS INbox construct. Respectively, these will be referred to as INbox^{W845G} and INbox^{F881A}. A double mutant containing both the W845G and F881A mutations, hereby referred to as INbox^{dbl}, was also generated and assayed

None of these constructs, regardless of whether or not they were processing-competent, caused a significant increase ANI in the cell-based CPC function assay (see figures 4.3.1 and 4.3.2 for representative micrographs, and panel A of Figure 4.4.1 for quantifications). This would indicate that CPC function is not disturbed under these conditions. Previously published reports have already indicated that the region encompassing W845 is required for Aurora B binding (Honda et al., 2003; Sessa et al., 2005; Xu et al., 2009). However, in Honda *et al.*'s study, a truncated INbox fragment lacking the region surrounding F881 could efficiently pull down Aurora B, albeit inactive, from cell lysates when these were co-expressed in Sf-9 cells. The results I have obtained suggest that both W845 and F881 are essential, not only for the activation of Aurora B, but also for INCENP binding to the kinase.

4.4 THE TSS MOTIF IS DISPENSIBLE FOR THE DOMINANT NEGATIVE EFFECT OF THE SOLUBLE INBOX

Another highly, if not absolutely, conserved (see Figure 5.2.1) INbox feature is the TSS motif present within its distal end. Although not involved in Aurora B binding, this motif is phosphorylated by Aurora B and is required for the full activation of the kinase. Phospho null and mimic INCENP mutants cannot rescue the loss of endogenous INCENP despite only modest drops in Aurora B activity as measured by histone H3 serine 10 phosphorylation (Xu et al., 2010).

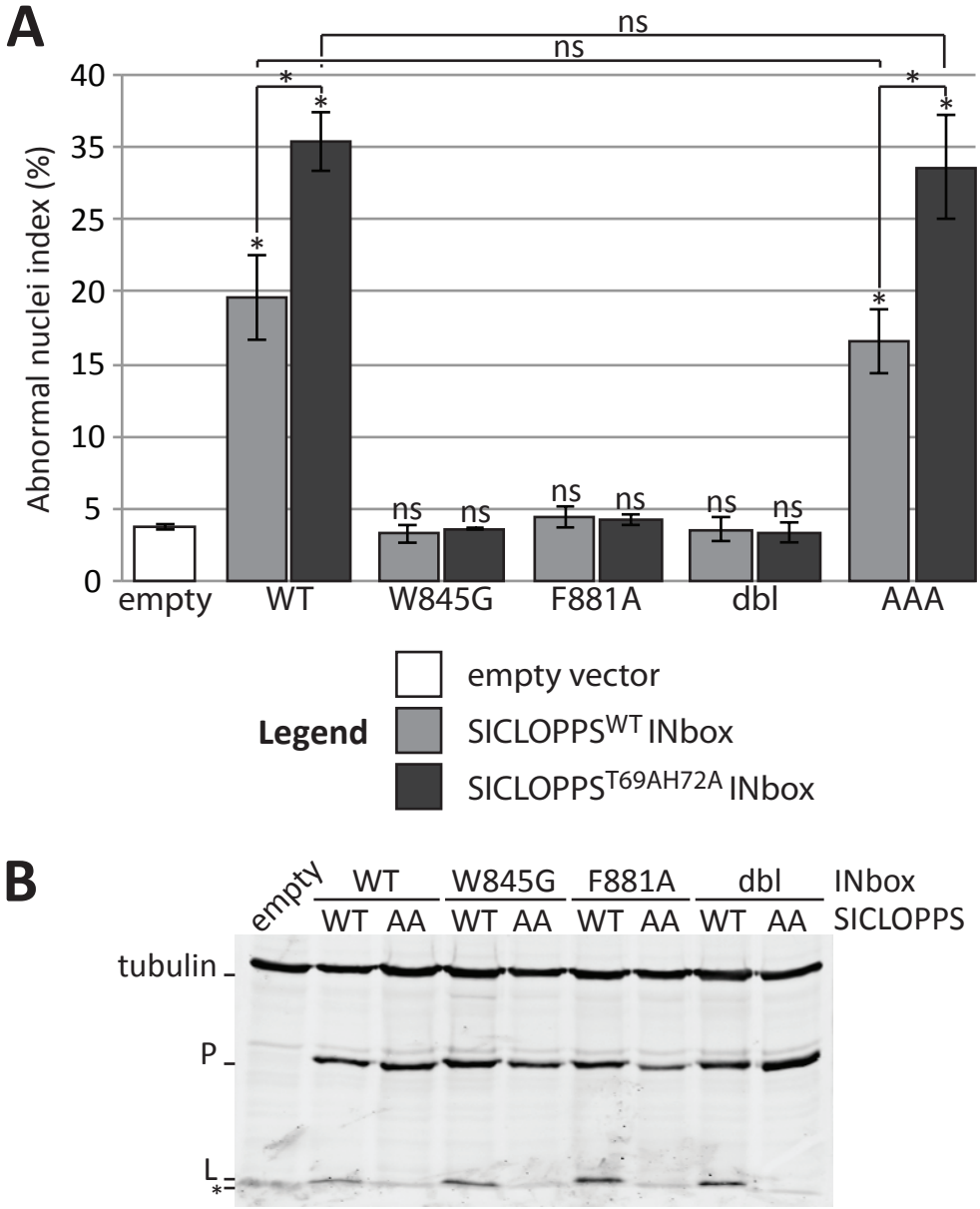


Figure 4.4.1 - Effect of the expression of the whole INCENP INbox or mutants thereof in HeLa cells. Constructs were transiently expressed for 48 hours, the last 24 hours of which under puromycin selection. (A) Quantification of the abnormal nuclei index of HeLa cells transiently transfected with INbox SICLOPPS constructs. The statistical significance of experimental samples compared to empty vector controls is indicated immediately over the error bars (Mann-Whitney U test, *=p>0.05, ns=not significant; n=3, >1000 cells scored per replicate; error bars= +/-SEM). (B) Processing of the SICLOPPS INbox constructs as detected by western blotting with anti-myc antibody and anti-tubulin used as a loading control. The precursor (P) and linear product (L) bands are as defined in Figure 3.2.1; the asterisk marks a non-specific band.

A pair of SICLOPPS and SICLOPPS^{T69AH72A} constructs was made in which all residues within the TSS motif are substituted with alanine (hereby referred to as INbox^{AAA}). In contrast to the W845 and F881 residue mutants, both INbox^{AAA} constructs elicited a strong increase in ANI comparable to their respective splicing proficient/deficient INbox^{WT} constructs (Figure 4.3.2 and Figure 4.4.1, panel A). This result suggests that the TSS motif is dispensable for Aurora B binding *in vivo* and is consistent with Honda *et al*.'s *in vitro* experiments (Honda et al., 2003). As with the INbox^{WT} constructs, the decrease in ANI between the SICLOPPS^{T69AH72A} and SICLOPPS versions of the constructs was comparable to the extent of processing (panel A of Figure 4.4.1 and data not shown).

4.5 EXPRESSION OF SOLUBLE INBOX CAUSES AURORA B MISLOCALISATION

INCENP is known to be required for both Aurora B activation and its localization during mitosis. To test whether the soluble INbox affected Aurora B localisation, the effects of transient expression of SICLOPPS^{T69AH72A} INbox^{WT} and INbox^{dbl} in mitotic HeLa cells were examined by immunofluorescence microscopy. The SICLOPPS^{T69AH72A} versions of these constructs were selected for this assay because the INbox^{WT} construct was most potent in its processing deficient form. The INbox^{dbl} construct was used because it bore both the W845G and F881A mutations that nullified the dominant negative effect of the soluble INbox. Cells transfected with an empty vector served as a negative control.

Consistent with the hypothesis that the soluble INbox might interfere with the interaction between INCENP and Aurora B, Aurora B was frequently mislocalised in mitotic cells expressing SICLOPPS^{T69AH72A} INbox^{WT} (Figure

4.5.1). Compared to control cells transfected with an empty vector, the localisation of Aurora B was significantly affected in all stages of mitosis in cells transfected with INbox^{WT} but not INbox^{dbl} ($p<0.001$ and $p>0.05$ respectively, n=3, Fisher's exact test; Figure 4.5.2, panel A). In the presence of INbox^{WT}, Aurora B was either completely or partially mislocalised in the vast majority of cells fixed in prophase to anaphase (Figure 4.5.2, panel A). The localization of Aurora B in cells during telophase and cytokinesis was less severely affected by INbox^{WT} expression and normal localization could be observed in roughly half the cells scored.

Given the CPC's role in several processes during mitosis and cytokinesis, mitotic indexes and progression were also monitored in these samples. All samples had comparable mitotic indexes (empty vector: 4.3% +/- 0.6, INbox^{WT} 4.3% +/-0.7, INbox^{dbl} 4.0% +/-1.5). However, samples transfected with SICLOPPS^{T69AH72A} INbox^{WT} exhibited a significant increase in prometaphase cells and relative to the empty vector control ($p<0.001$, n=3, Chi square test; Figure 4.5.2, panel B). The increase in the proportion of cells in prometaphase suggests that the soluble INbox might cause a chromosome alignment, but not mitotic checkpoint, defect. An appreciable drop in the proportion of cells present in cytokinesis was observed in samples expressing INbox^{WT} but this was not statistically significant.

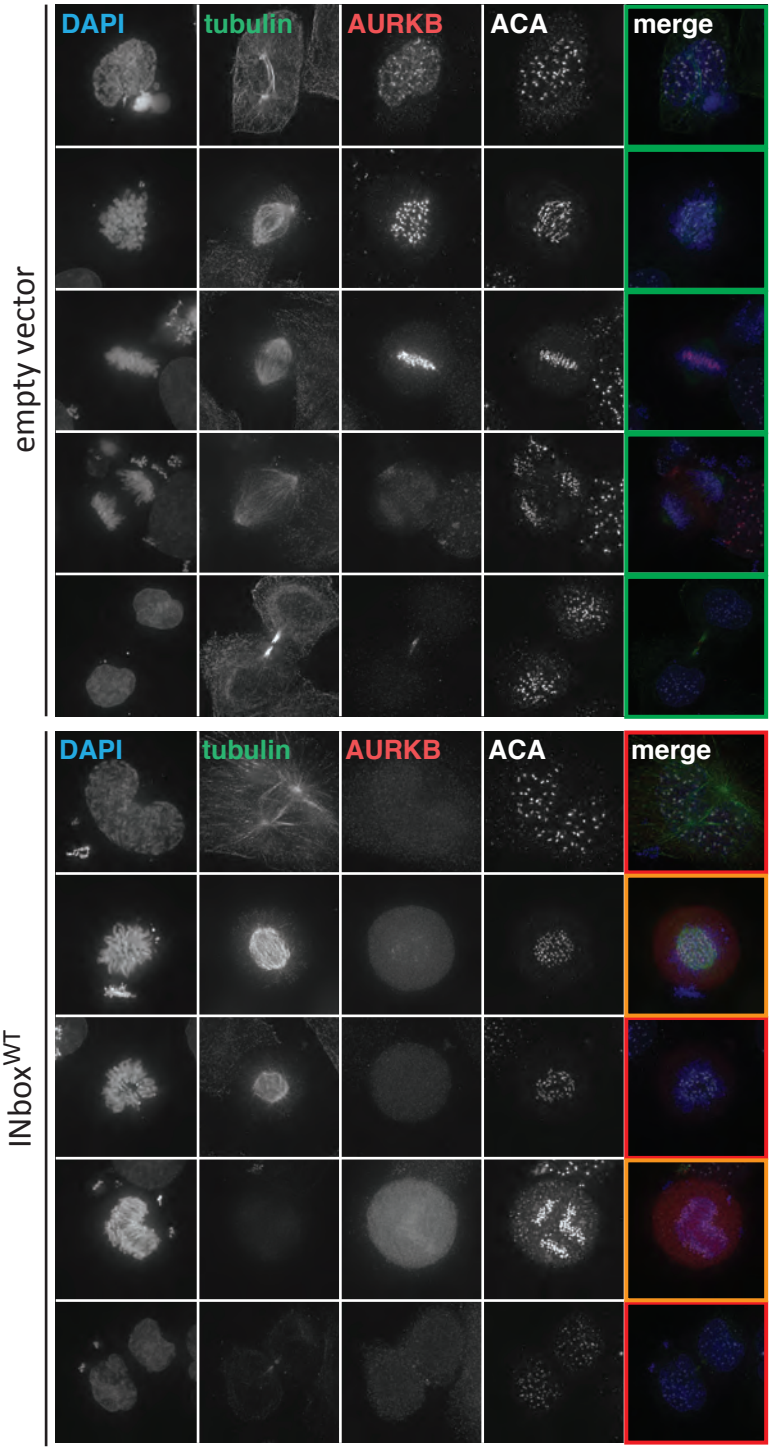


Figure 4.5.1 - Localisation of Aurora B during mitosis in puromycin-selected transiently transfected cells after 48 hours as observed with a 100X objective. Cells were either transfected with an empty vector or with one expressing SICLOPPS^{T69AH72A} INbox^{WT}. Cells were counterstained with anti tubulin and anti centromere antibody (ACA) to visulase the mitotic spindle and centromeres, respectively. Coloured boxes around the merged panels indicates how these representative cells would be classified according to the scoring criteria employed in Figure 4.5.2.

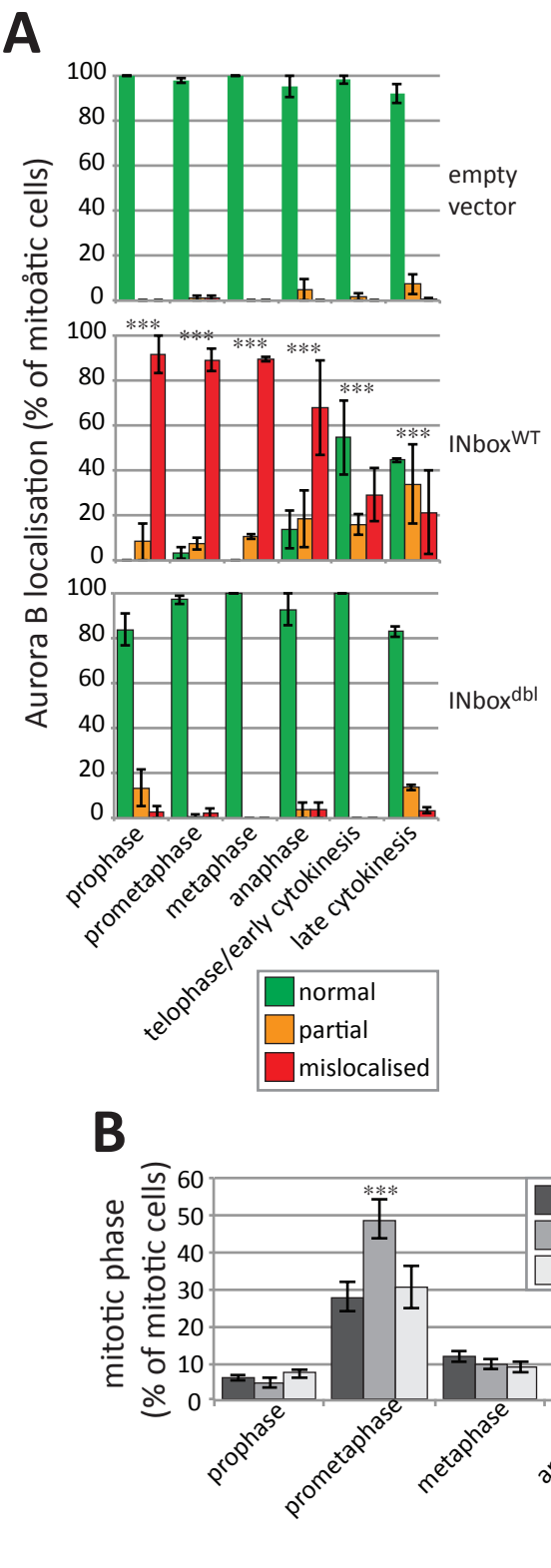


Figure 4.5.2 - (A) Quantifiation of the localisation of Aurora B in puromycin-selected cells transiently expressing SICLOPPS^{AA} INbox^{WT}, INbox^{dbl} or an empty vector control for 48 hours. The distribution of normal, partially and mislocalised cells was compared for each phase were compared to those found in the empty vector control sample using Fisher's exact test. (B) Proportion of the mitotic cells from the same samples present in each stage of mitosis and cytokesis. The proportion of cells in each stage were compared to that of the empty vector control using the Chi square test. In both A and B: n=3; >100 cells per replicate; error bars: +/- SEM; *** = p<0.0001; unmarked samples are not significantly different from the control.

4.6 DISCUSSION

The results presented in this chapter offer a preliminary insight into the consequences of abrogating the Aurora B-INCENP interaction. I will first reflect on the significance and caveats of these results. Following this, I will consider the performance of SICLOPPS-based constructs in this context.

As cautioned earlier (section 3.5), ANI increases cannot necessarily be directly attributed to impaired CPC function. This is because other proteins and complexes also play important roles in both chromosome segregation and cytokinesis. It is, however, highly probable that the increase in ANI elicited by $INbox^{WT}$ and $INbox^{AAA}$ constructs observed in HeLa cells is due to CPC function impairment as these soluble constructs consist of the Aurora B binding domain of INCENP. Both $INbox^{WT}$ and $INbox^{AAA}$ presumably bind to the N-terminal lobe of Aurora B, thereby occluding INCENP and the linked targeting module. This is further supported by the mislocalisation of Aurora B in mitotic cells expressing $INbox^{WT}$.

Expression of INbox constructs lacking either the W845 or F881 conserved residues, or both, fails to impair CPC function. This is reflected by the unchanged ANI in samples expressing mutant constructs relative to empty vector controls. This is interpreted as resulting from a lack of mutant INbox binding to Aurora B. This would indicate that the W845 and F881 residues are each essential not only for Aurora B activation, but for INCENP binding to the kinase's N-terminal lobe.

Further to its effect on CPC function as observed with the phenotypic CPC function assay, overexpression of $INbox^{WT}$ causes mislocalisation of Aurora B and an accumulation of cells in prometaphase. The general mislocalisation of Aurora B supports the hypothesis that the dominant

negative effect of the soluble INbox is due to its interfering with the Aurora B-INCENP interaction. We observed that the extent of Aurora B mislocalisation was greatest in cells fixed in the earlier stages of mitosis. The appreciable drop in the severity of mislocalisation in the later phases could be due to the use of transiently transfected constructs. That is, cells with lower soluble INbox levels, which would occur due to expression heterogeneity, may be less severely affected in both CPC localization and mitotic progression. This may also explain our finding that the slight change in the proportion of cells present in late cytokinesis was not significant despite the fact that CPC function is required for this process.

The dominant negative effect of overexpressing the INbox domain was less pronounced than that of overexpressing the kinases dead mutant of Aurora B. This could be due to differences in expression levels, different mechanisms of action, or a mixture of the two. In both instances, endogenous Aurora B, which possesses basal activity, should be present in the cells. However, kinase dead Aurora B might not only obstruct the binding of endogenous Aurora B to INCENP, but also to its substrates.

Despite being flanked by the native extein residues, the processing of the INbox SICLOPPS constructs was lower than that observed for the test construct (contrast Figure 3.3.2 and panel B of Figure 4.4.2). It is possible that the accessory residues employed here do not promote intein processing as efficiently as those present at the same positions in the SICLOPPS test construct (see Appendix I). It is possible that the physical properties of the extein (e.g. its length and flexibility/rigidity) may also affect intein processing. The SICLOPPS test construct encoded a 14 residue CP in contrast to the 75 residue long INbox CP. How best to promote the circularization of proteins

and peptides with different sequences and properties remains to be ascertained.

When comparing the respective ANI increases elicited SICLOPPS and SICLOPPS^{T69AH72A} versions of the INbox^{WT} and INbox^{AAA} constructs, it appeared that the dominant negative effect was elicited primarily by the linear precursor, not the CP product. If this interpretation is correct, there are several possibilities why binding of these INbox constructs may have been impeded by intein processing. Firstly, and most likely, the constraint imposed by circularization may have 'locked' the INbox peptides in a conformation that is not conducive to binding. This unfavourable conformation could have been local or global, affecting a key stretch of residues or preventing the whole peptide from crowning the Aurora B N-terminal lobe. Another minor possibility, which should nevertheless be considered, is that intein processing somehow might promote the loss of the INCENP peptide rather than its stabilisation. A study monitoring circularization of whole GFP by SICLOPPS highlighted that, further to the circularized GFP, some linear GFP was also produced (Zhao et al., 2010). Should the linear INCENP fragment in isolation (i.e. without the flanking inteins) be produced to a significant extent and be unstable in cells, SICLOPPS processing would diminish the overall amount of INbox peptide in the cell.

Blocking the Aurora B-INbox interaction resulted in cell division defects that are similar, if milder, than those of overexpressing kinase-dead Aurora B. As demonstrated using the chicken DT40 INCENP knockout, even a mild drop in CPC function is incompatible with cell proliferation and survival (Xu et al., 2010). This suggests that disrupting the Aurora B-INbox interaction will eventually lead to cell death. Whether this is indeed the case should be

verified by assaying the effects of soluble INbox overexpression over longer time periods.

Furthermore, the localization of other CPC components and the activation status of Aurora B in the presence of the $INbox^{WT}$ and $INbox^{AAA}$ constructs would be highly informative and remain to be verified. These and other experiments will be discussed further in the Perspectives section.

CHAPTER 5: SCREENING A RATIONALLY-DESIGNED CYCLIC PEPTIDE INBOX MIMIC LIBRARY

5.1 INTRODUCTION

The results presented in the previous chapter provide additional evidence towards the validation of the Aurora B-INbox interface as an anti-mitotic target. However, it is a large interface and our efforts to dissociate it employed a construct with a 75 amino acid long stretch of homology. It was unclear whether the Aurora B-INbox interface is a target that can be modulated by smaller compounds, which would indicate that it might be druggable. I thus proceeded to express smaller pieces of the INbox *in vivo* to attempt to identify fragments capable of recapitulating the effects of the whole INbox^{WT} and INbox^{AAA} constructs.

In order to better understand INbox library design features and the first-pass screen presented within this chapter, the chronological order of events must first be clarified: both these sections of the project were completed early on during piloting. This was immediately following the verification of SICLOPPS test construct processing in HeLa cells (section 3.3). Thus, results showing that the soluble INbox was more potent in processing deficient than competent forms were not available at the time. At the stage of library design, the inhibition of Aurora B activity by soluble INbox was predicted, if unconfirmed.

Therefore, the ultimate version of the cell-based CPC functional assay presented in chapter 3 and employed in chapter 4 was informed by this early work rather than *vice versa*. As the second-pass screen also presented later in this chapter employed the same improved protocol used in generating the results presented in earlier chapters, these series of experiments remain

comparable and the controls, first shown in sections 3.3 and 3.4, relevant. The rationales laid out in sections 5.2 to 5.4 reflect my thinking *at the time* these experiments were devised and carried out. The discussion section of this chapter contains a reassessment of the library design and first pass screening protocol in light of subsequently obtained results.

5.2 LIBRARY DESIGN

In contrast to the globular N-terminal lobe of Aurora B, both the *Xenopus* and human INCENP INbox adopts an extended conformation containing interspersed α-helical stretches (Elkins, 2012; Sessa et al., 2005; also recall Figure 1.3.3). The intention was to design a library of fragments matching the Aurora B – INbox interface based on primary amino acid sequence. In contrast to the INbox, some adjacent residues within the Aurora B side of the interface are non-contiguous, which would have necessitated a more complex mimicry design strategy. Thus I focused on INCENP rather than Aurora B sequences.

Rather that base library design on the full length INbox, I focused on a smaller fragment of the INbox spanning residues 834 to 894 (Figure 5.2.1.A). The rationale for this is that this region corresponds to the one used by Sessa and colleagues in their structural study of the *Xenopus* Aurora B-INbox complex (Sessa et al., 2005). Thus, it is known to be associated with the Aurora B N-terminal lobe. In particular, residues 842 to 885 are visible in the resulting crystal structure (Figure 5.2.1.B).

Previous screens of SICLOPPS combinatorial libraries utilized cylic peptides (CPs) ranging in size from penta- to octapeptides (Cheng et al., 2007; Kinsella et al., 2002). A CP length of 8 – an invariable cysteine plus 7 variable residues – was arbitrarily selected for the core library (Figure

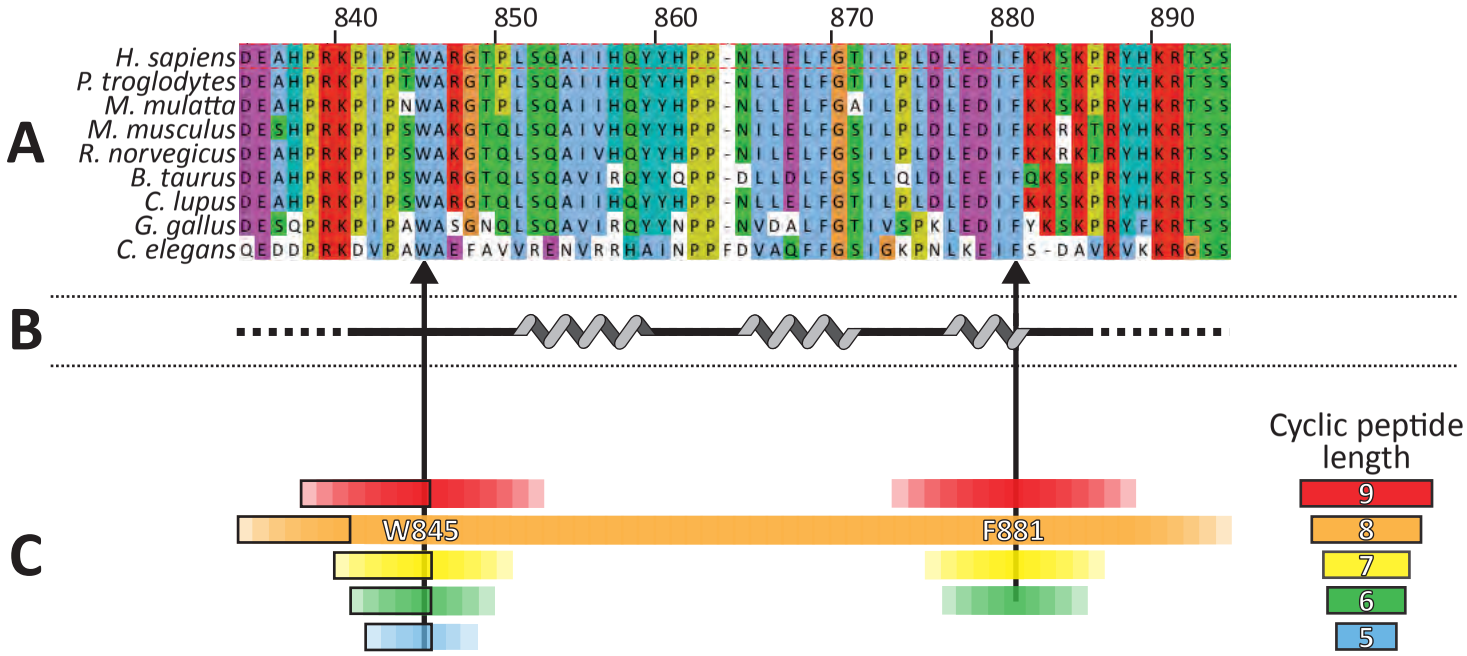


Figure 5.2.1 - Overview of the INCENP INbox library design rationale. All panels are aligned so that the contents of panels B & C map to the residues vertically above them in panel A. (A) The core INbox fragment is highly conserved and residues W845 and F881 (arrows) are absolutely conserved across all homologs. hINCENP residue numbers are indicated above the alignment. (B) Secondary structure of the INCENP core fragment in Sessa et al.'s study of the *Xenopus* Aurora B/INbox complex. Dashed lines indicate residues not visible in the crystal structure. (C) Coverage of the core library (orange) and sublibraries of different lengths (red, yellow, green and blue). The cyclic peptide length indicated includes the invariable cystein residue present in extein position +1. For example, length of 5 would indicate a cyclic peptide consisting of a cystein and four residues homologous to the INbox.

5.2.1.C). Optimal cyclic peptide backbone length will vary depending on the epitope and cannot be predicted given currently available means. Therefore, additional sub-libraries centered on residues W845 and F881 – two conserved residues critical for Aurora B binding and activation – were generated with varying CP backbone lengths (Figure 5.2.1.C). Totaling 97 members, codon-optimized reverse primers encoding the library were manually designed using the sequence of the human INbox. All libraries consist of overlapping peptides with a single residue shift between adjacent library members. The full CP library and corresponding primer sequences can be found in Appendix II.

In designing primers, a sequence of two lysine residues was accidentally changed to a single one in the F881 5 variable residue sub-library. This mistake affects 4 library members (A11-B2 according to naming system used in Appendix II)

5.3 VECTOR OPTIMIZATION AND LIBRARY CONSTRUCTION

From my earlier attempts at expressing SICLOPPS constructs with and without selection (section 3.3), it was evident that any screen would require the selection of transfected cells in order to minimize background noise from untransfected cells. I therefore optimized the pIRES puro2 vector to make it compatible with library construction.

Two restriction sites lie sufficiently close to the variable CP-encoding region to be utilised in library construction: *NheI* and *AflII*. However, both sites are also present in the pIRES puro2 multiple cloning site (MCS). As *NheI* digest results in a high level of vector re-ligation (R. Almeida, personal communications), *AflII* was favoured and its recognition site within the MCS was removed by site-directed mutagenesis. The modified MCS was then

substituted into vectors containing both the SICLOPPS^{WT} and SICLOPPS^{T69AH72A} versions of the test constructs employed in section 3.3.

Concurrently with this, in the hopes of saving time, I sought to have the library built commercially and obtained quotations from DNA2.0 and GeneArt. Disappointingly, the quotes were in the order of US$20,000 – a cost deemed prohibitive. I therefore adapted the method outlined by Abel-Santos and colleagues (Abel-Santos et al., 2003) to build the library manually, simplifying it for use with discrete oligos rather than pooled oligos. The abridged protocol omits the zipper PCR step as, in this instance, maximizing cloning efficiency is a lesser concern than in the generation of combinatorial libraries.

The template used for PCR amplification of library sequences was an *NcoI*-digested plasmid bearing the SICLOPPS test construct. The *NcoI* site lies upstream from the initiation codon, but downstream of the *BamHI* site used during library construction; the forward primer used to amplify library inserts has an extension containing the *BamHI* site. This ensures that only the digested PCR products will contain a *BamHI* sticky end, minimizing template carry-over and obviating the need to gel extract the PCR product. In lieu of gel extraction, a kit-based PCR cleanup was used. A detailed library generation protocol can be found in the methods section.

94 of the 97 library members designed were generated. Those still missing are: C4, C11 and E2 (refer to Appendix II for further details). Library member C4 lies in the W845 region of the 7 variable residue core library whereas C11 and E2 lie in the F881 region of the 7 and 8 variable residue libraries, respectively. The sequence coverage of the completed library is shown in Figure 5.3.1.

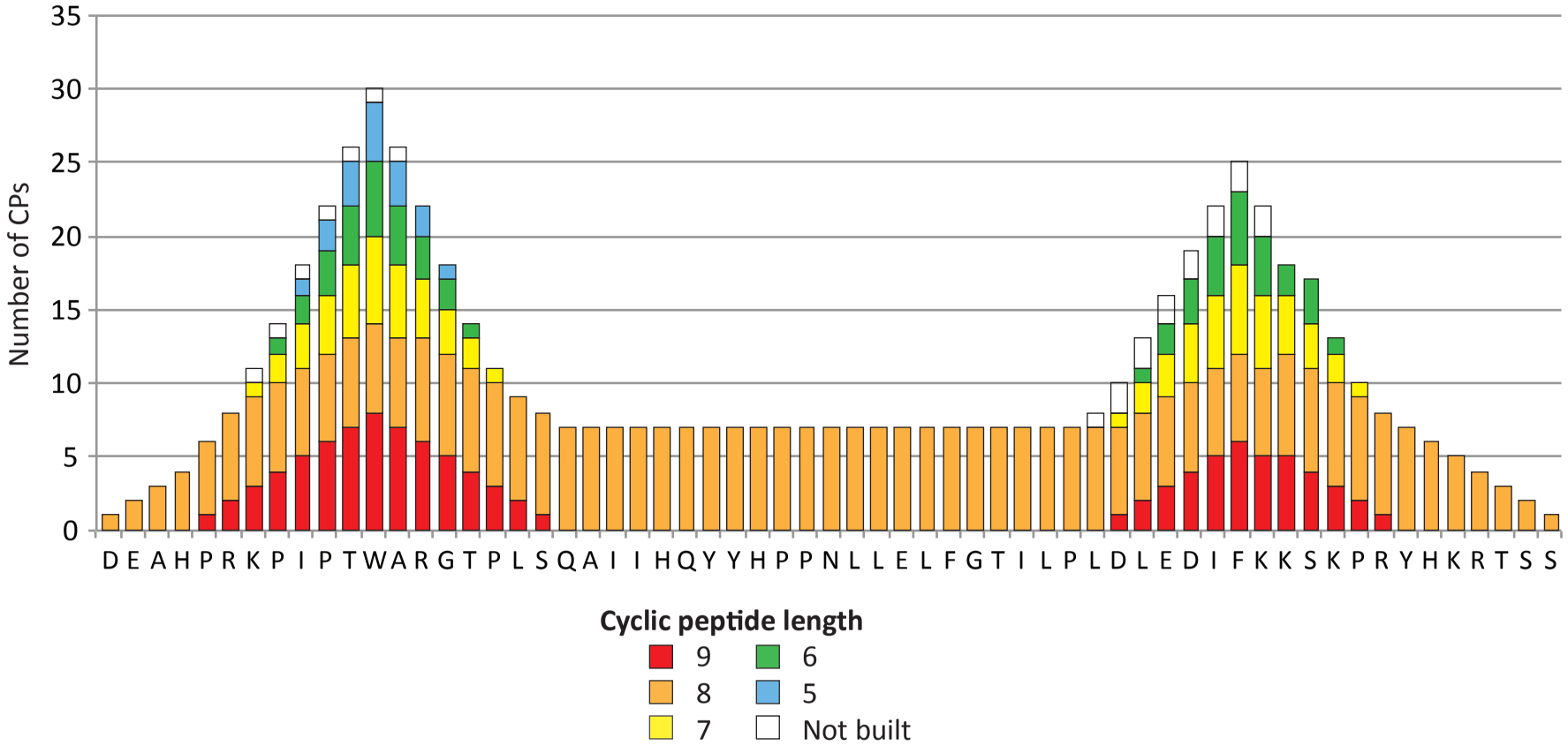


Figure 5.3.1 - Coverage of the completed INCENP INbox library with regards to the primary amino acid sequence. Colours indicate the overall length of cyclic peptide (CP) constructs, which all contain an invariable cystein residue in position +1 followed by a variable length of residues with direct homology to the INbox.

5.4 LIBRARY SCREEN – FIRST PASS: ABNORMAL NUCLEI INDEX

The first pass screen protocol was a precursor to the CPC function assay. The rationale for readout selection and screen design and has already been laid out in section 3.3.

The first pass screen was carried out during piloting and 97 INbox library members were to be assayed initially. For these reasons it was desirable to streamline the assay protocol so that it could be completed rapidly. Each library member was assayed in duplicate and the ANI was scored from 5 random fields of view as visualized with a 40X objective. Thus, the amount of replicates and cells scored was lower than in the final version of the CPC function assay. The triplicate scoring of >1000 cells was resumed in the final version of the CPC function assay during which fewer library members needed to be assayed. Another notable difference between the two protocols is the use of poly-L-lysine coated, rather than uncoated, coverslips in the initial protocol. Initially, cells were simply DAPI stained. Phalloidin staining was rapidly employed to highlight cellular periphery and improve scoring reliability.

To determine the rate of increase in the frequency of abnormal nuclei, abnormal nuclei indexes (ANI) measured for each coverslip were normalised to the average ANI measured from a pair control coverslips transfected with empty pIRES puro2 vector in parallel to the experimental samples. Thus, we compared ANI fold change rather than raw ANI. This was done in order to attenuate the effects of any variability in ANI due to passage number or cell confluency prior to seeding. The empty vector control was also used to determine the spread in abnormal nuclei frequency and establish hit thresholds. The ANI in the empty vector control population over the course of screening was normally distributed (Shapiro-Wilk test, $p=0.89$) and a hit

threshold was set at an increase in abnormal nuclei frequency >2 standard deviations from the mean background ANI – this was calculated to be 1.35X but arbitrarily rounded up to 1.4X to increase screen stringency. The ANI increase typical of Aurora B overexpression controls can be found in panel A of Figure 5.5.1 in the next section.

Of the 94 library members assessed in duplicate, 11 library constructs – hereby referred to as strong initial hits – elicited an ANI increase greater than the cutoff in both replicates. A further 9 library members were deemed moderate initial hits as their average ANI increase – as opposed to both individual replicate – was greater than 2 standard deviations from the control. In contrast to this, no library members showed an average decrease in abnormal nuclei frequency greater than 2 standard deviations from the same control. The sequence of hit constructs is presented in table 5.4. Figure 5.4.1 shows the proportion of hit constructs relative to the number of library members screened. Peptide sequences corresponding to 80% of the strong hits and 55% of the moderate hits identified during this initial screen were homologous to regions containing either INCENP W845 or F881. Further to this, another third of the moderate hits lay in regions immediately proximal or distal to these two critical residues (Regions highlighted in green and red in Table 5.4 and subsequent figures). Of the remaining 3 hits, 2 possessed sequences that did not overlap with that of any other hit.

Hits were concentrated in two regions: one centered on INCENP W845 and another encompassing F881 and residues distal to it (Figure 5.4.2, uppermost heat map). These two regions also overlap with the regions of highest library coverage (recall figure 5.3.1). However, the clustering of hits near INCENP W845 and F881 remains apparent even after normalizing the

INCENP residues	Fold ANI increase	Length
844-851	1.9	9
839-845	1.9	8
845.852	1.8	9
845-850	1.7	7
844-847	1.7	5
840-845	1.6	7
836-842	1.6	8
838-844	1.5	8
842-845	1.5	5
841-847	1.5	8
845-849	1.4	6
862-868	1.6	8
869-875	1.4	8
878-885	2.5	9
877-884	2	9
881-885	1.7	6
881-886	1.5	7
880-886	1.5	8
888-894	1.5	8
883-889	1.4	8

Table 5.4 - Sequence of the 20 hits derived from the intial screen of the INbox library. Hits are clusters according to the region of the INbox they map to (green: W845; blue: centre; red: F881) then ordered in descending order activity as indicated by abnormal nuclei index (ANI) increase, that is by the increase in ANI relative to an empty vector control transfected and scored in parallel. Bright and muted colours indicates strong and moderate initial hits, respectively, as defined in the text.

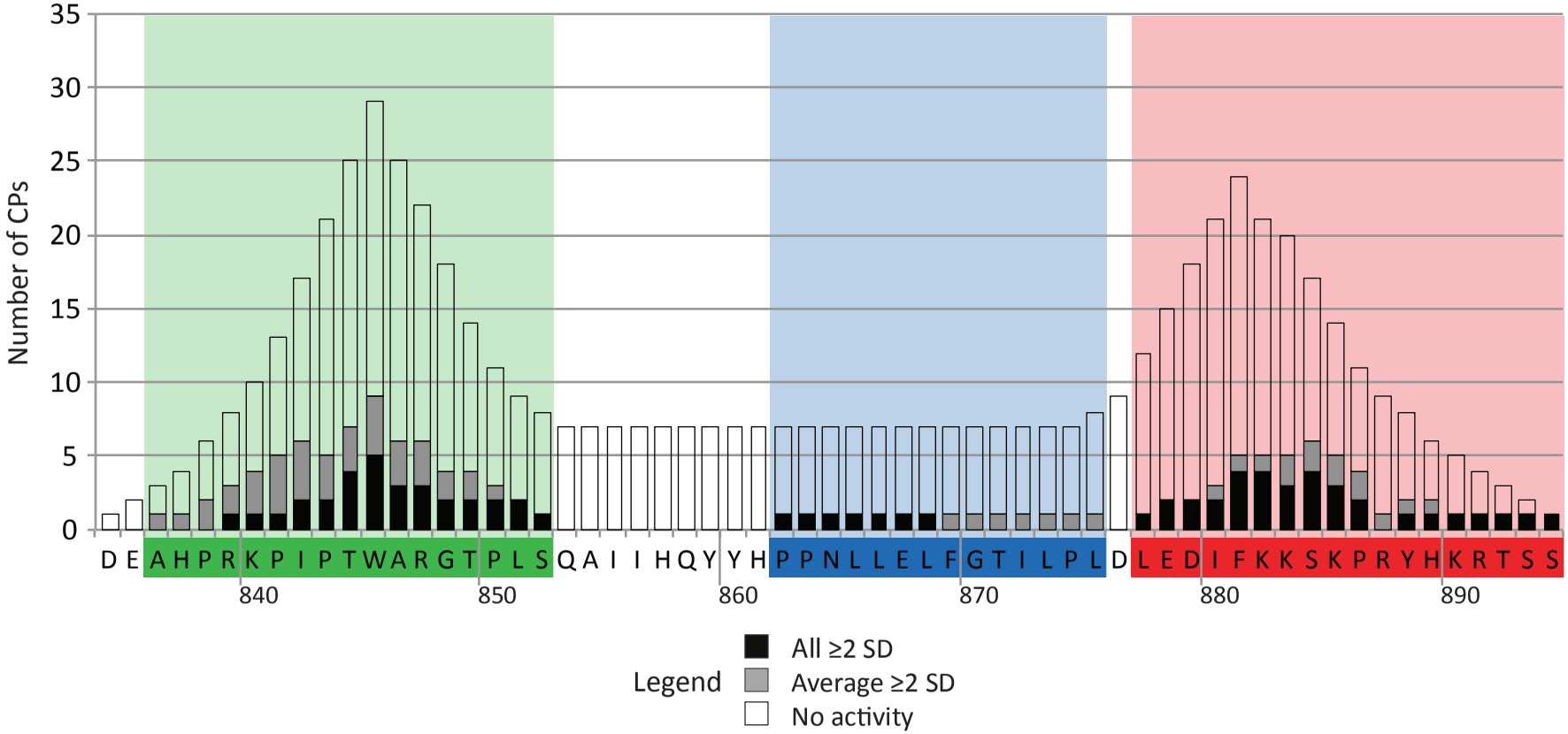


Figure 5.4.1 - Distribution of hit constructs sequences relative to the human INCENP INbox primary amino acid sequence. The colour of the bars indicates the activity of constructs mapping onto a given residue with respects to the >2 standard deviation (SD) threshold. Colour blocking indicate the hit regions outlined in Table 5.4.

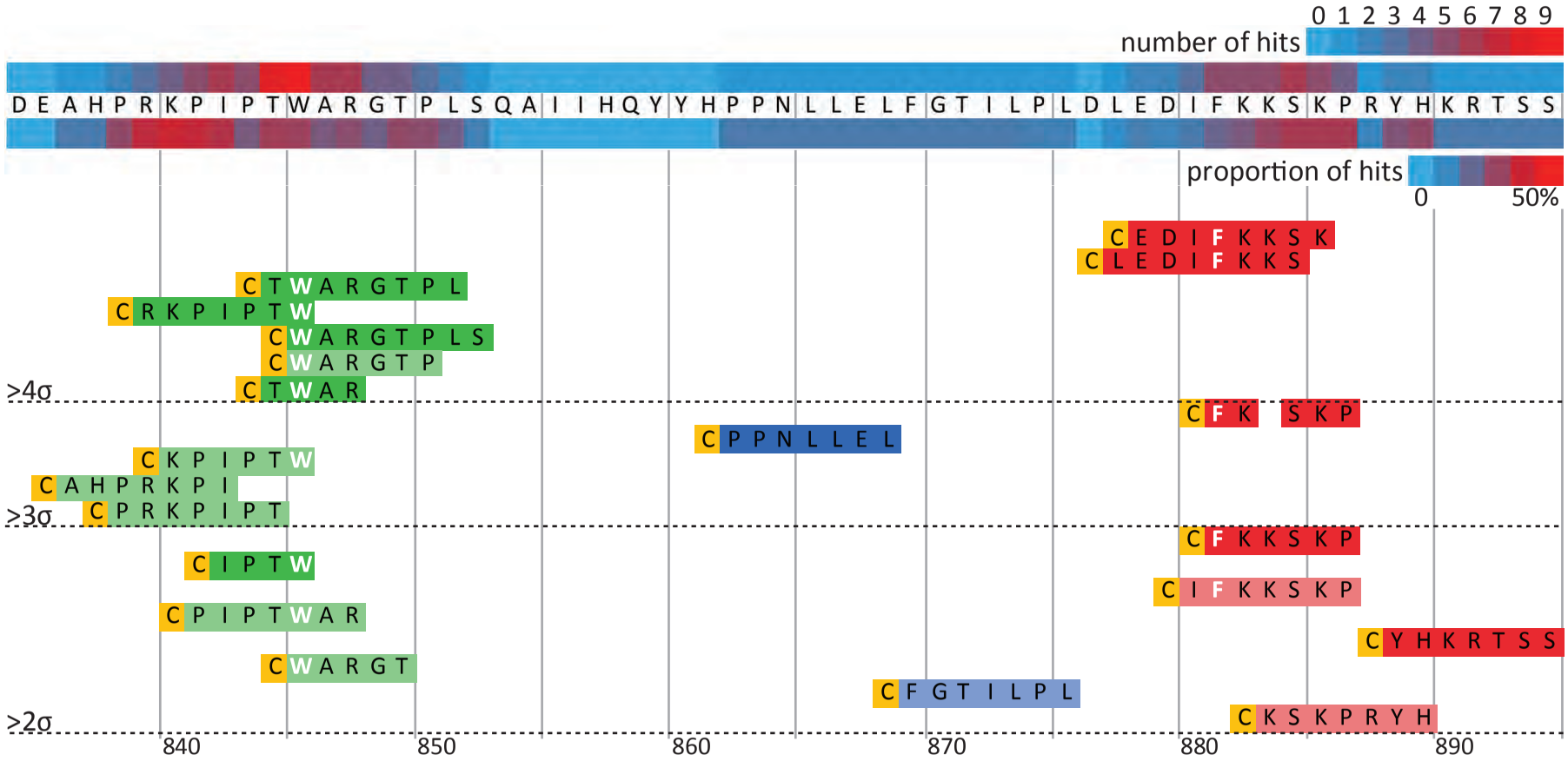


Figure 5.4.2 - Distribution of first pass screen hits with respect to the INbox primary amino acid sequence. Heatmaps indicate the number of hit constructs mapping to a given residue (top) or the proportion of hits relative to the total number of library consructs encompassing said residue (bottom). Hit construct sequences, ranked in descending order of ANI increase, are displayed directly beneath the corresponding residues and the dotted lines indicate standard deviation thresholds. Colour blocking corresponds with the hit regions outlined in Figure 5.4.1 and Table 5.4.

number of hit peptides to the number of library members (Figure 5.4.2, lower heat map).

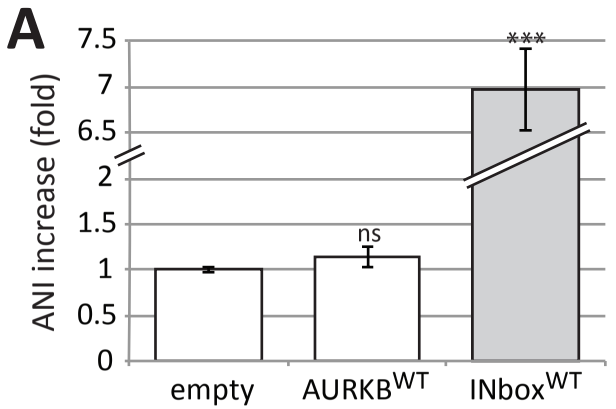
5.5 LIBRARY SCREEN – SECOND PASS: CYCLIC VS. LINEAR PRECURSOR

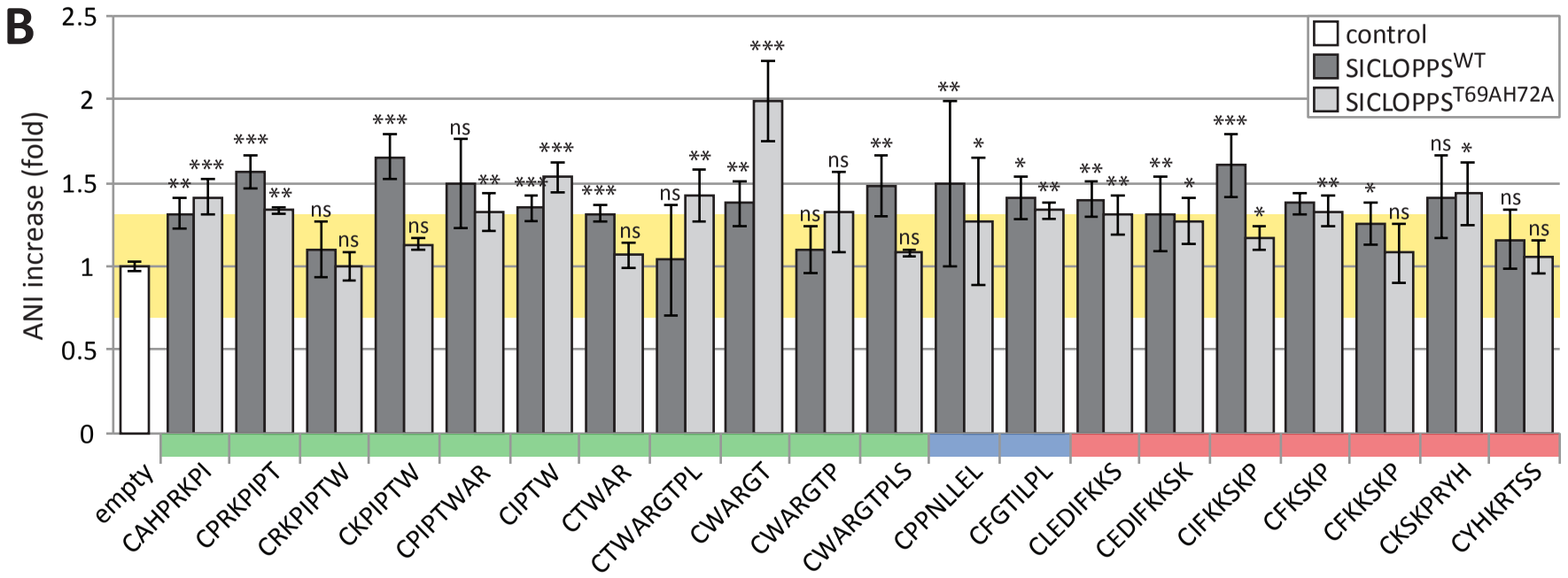
The initial INbox library screen yielded 20 hits from 97 constructs. To validate these hits, it was first necessary to verify if their effect was reproducible. Secondly, I wished to verify if constructs' activity was due to the cyclic peptide being produced rather than the split intein-flanked precursor. The reason for this was that we wished to ascertain whether the Aurora B-INbox interaction was amenable to being dissociated by small compounds. To ascertain this, SICLOPPS^{T69AH72A} versions of each hit construct were generated and assayed in parallel with their processing-competent counterparts. These mutants cannot be cyclized.

As previously indicated, methodological optimisation carried out following the first round of screening led to an increase in the number of replicates and cells scored. Because the use of poly-L-lysine coated coverslips impeded the detachment of dead cells following puromycin selection, uncoated coverslips were thenceforth utilised. The seeding protocol was revised to achieve an appropriate cell density for transient transfection and scoring.

The results of the rescreening are presented in figure 5.5.1 and summarized in table 5.5.1. Upon rescreening with the more stringent scoring protocol, 70% of the initial hits (14 of 20 hits) were reproducible but only 29% of these (4 of 14 reproducible hits) required SICLOPPS processing for their activity.

Figure 5.5.1 - (A) Abnormal nuclei index (ANI) increase elicited by control constructs in the CPC function assay carried out as described earlier. The ANI increase is a measure of a sample's ANI divided by the average ANI of a pair empty vector control carried out in parallel. (B) Second pass screening of initial screen hits alongside processing-deficient mutants thereof. The level of statistical significance between the sample and empty vector control is indicated above the error bars (error bars: +/- SEM; ns: non significant, *: p<0.05, **: p<0.01, ***: p<0.001; two-tailed Mann Whitney U test, n≥3, >1000 cells counted per biological replicate)





INCENP residues	Reproducibility	Processing dependency
844-851 CTWARGTPL	no	N/A
839-845 CRKPIPTW	no	N/A
845-852 CWARGTPLS	yes	yes
845-850 CWARGTP	no	N/A
844-847 CTWAR	yes	yes
840-845 CKPIPTW	yes	yes
836-842 CAHPRKPI	yes	no
838-844 CPRKPIPT	yes	no
842-845 CIPTW	yes	no
841-847 CPIPTWAR	no	N/A
845-849 CWARGT	yes	no
862-868 CPPNLLEL	yes	no
869-875 CFGTILPL	yes	no
878-885 CEDIFKKSK	yes	no
877-884 CLEDIFKKS	yes	no
881-885 CFKSKP	yes	no
881-886 CFKKSKP	yes	yes
880-886 CIFKKSKP	yes	no
888-894 CYHKRTSS	no	N/A
883-889 CKSKPRYH	no	N/A

Table 5.5.1 - Reproducibility and processing dependency of initial hit peptides during the second pass screen. Hits are presented in the same order and with the same colour coding employed in table 5.4. Reproducibility indicates that the SICLOPPS^{WT} version of the construct elicited a satistically significant ($p<0.05$) increase in ANI relative to the empty vector control. Processing dependency indicates that only the CP producing construct elicited a significant increase, i.e. that the SICLOPPS^{T69AH72A} version of the construct did not have a significant effect on ANI increase. Constructs whose effects were not reproducible in the first instance could not be considered to be processing dependent, thus are marked ‘not applicable’ (N/A).

As already mentioned in the context of the full length INbox expression constructs in the previous chapter, split-intein mediated protein splicing efficiency can be affected by extein sequence. It was therefore expected that some of the library constructs would not produce CPs, or do so poorly. I wished to assess the degree of SICLOPPS processing throughout the library and verify whether there was a relationship between processing efficiency and activity. A random sample of hit and non-hit peptides were assayed in parallel for splicing efficiency and ANI increase (Figure 5.5.2.A and B). No correlation could be detected between SICLOPPS processing rates and peptide activity (Pearson's r=-0.11; p=0.54; n=35). This lack of correlation indicates that processing alone does not determine a construct's activity, suggesting that activity is sequence dependent.

Further to the lack of correlation, analysis of the processing efficiency throughout the library highlighted its heterogeneity (processing range= 1 to 99%, mean=36.1, SD=28.36). Unfortunately, this processing heterogeneity confounds direct comparison of the effects of SICLOPPS-generated CPs as comparable levels of these cannot be reliably expressed. Further to possessing activity relative to non-hits, the activity of any genuine CP hits should be concentration dependent. Genetically encoded constructs are ill-suited to assaying a wide concentration gradient. Thus, the SICLOPPS technology becomes limiting.

As both direct comparisons of hit CP activities and further characterization of these are logical continuations of this project, I sought to obtain chemically synthesized versions of these peptides. To prioritize the peptides to be synthesized, hits were ranked (Table 5.5.2). The ranking is semi-arbitrary: constructs with the highest ANI increase (as a measure of both mean ANI increase and SEM) were favoured and then ranked further

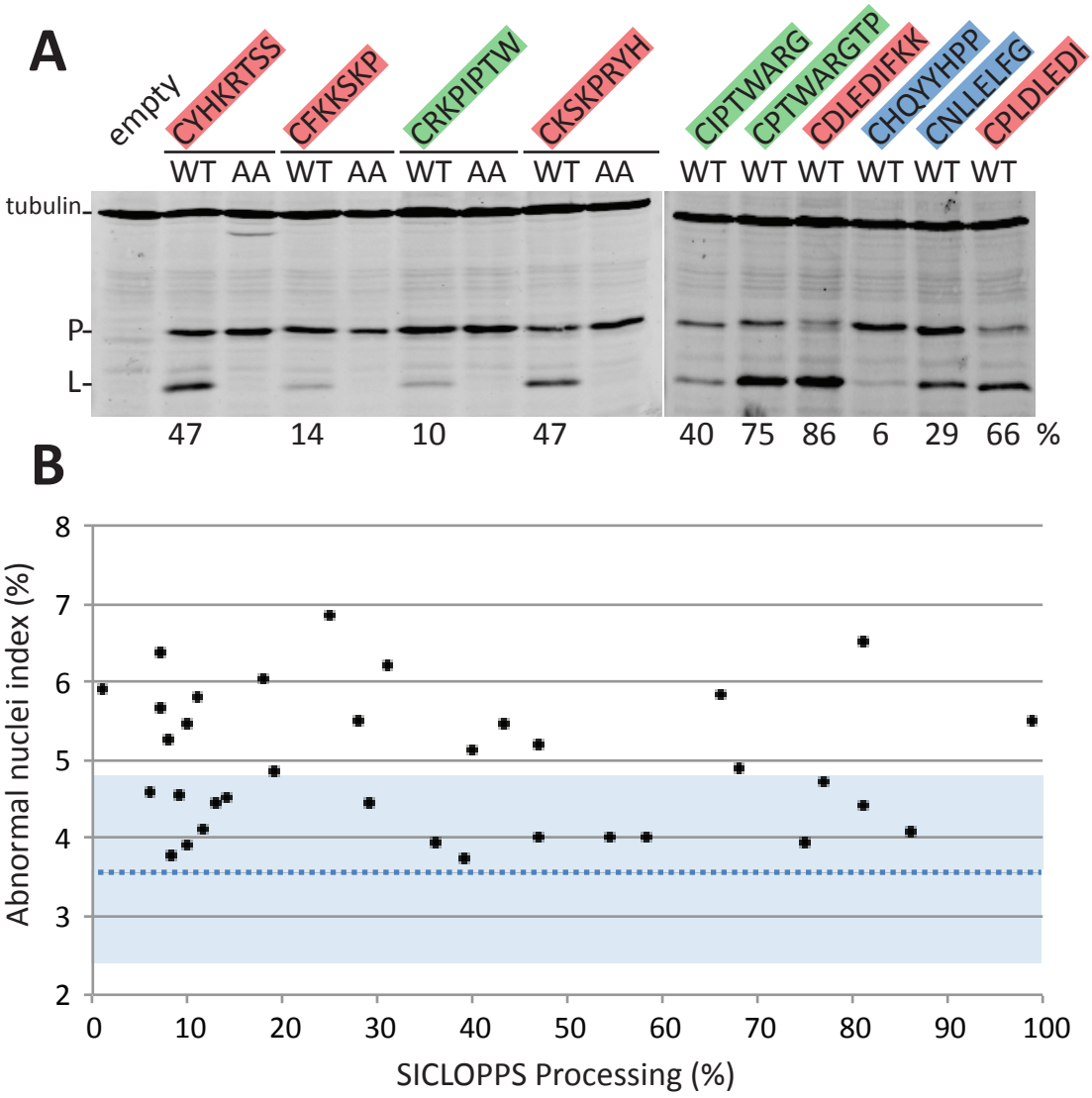


Figure 5.5.2 - A sample of randomly selected hit and non hit library constructs were assayed in parallel for SICLOPPS processing by western blot and ANI increase using the CPC function assay. (A) Representative quantitative western blot analyses of hit peptides in their processing-competent (WT) and deficient (AA) forms (left) and non hit peptides (right). Recall that processing deficient mutants were only prepared for constructs identified as initial hits in the first pass of the screen. The extent of processing (i.e. the intensity of 'L' band divided by the sum of the intensity of the 'L' and 'P' band) is indicatedas a percentage below each processing-competent lane. The precursor (P) and linear product (L) bands are as defined in figure 3.3.1. (B) No correlation between between SICLOPPS processing rates and abnormal nuclei index (ANI) could be detected (Pearson's r=-0.11; p=0.54; n=35). Mean background ANI as measured in control samples transfected with empty vector is indicated by a dashed line; the shaded region delineates +/- 2 standard deviations from the background ANI.

according to the effect of the SICLOPPS^{T69AH72A} mutants thereof. A number of constructs whose SICLOPPS^{T69AH72A} mutants elicited a small yet statistically significant increase in ANI relative to the empty vector control were also included in the prioritized list. From this list, a subset of peptides – and scambles thereof – representing the 3 hit regions identified during the initial library screen were synthesized by our collaborators Daniel St Cyr, Samir Bouayad-Gervais and Mike Tyers at the Université de Montréal. The planned use of these peptides will be discussed in the Perspectives section (Chapter 7).

		SICLOPPS^{WT}		SICLOPPS^{T69AH72A}			
Rank	Sequence	ANI increase	SEM	ANI increase	SEM	Synthesized	Scramble
1	CKPIPTW	1.7	0.14	1.1	0.03	yes	CPWITKP
2	CPRKPIPT	1.6	0.10	1.3	0.02	no	N/A
3	CIFKKSKP	1.6	0.19	1.2	0.08	yes	CKPFKSIK
4	CTWAR	1.3	0.05	1.1	0.08	no	N/A
5	CWARGTPLS	1.5	0.18	1.1	0.02	no	N/A
6	CPPNLLEL	1.5	0.50	1.3	0.38	yes	insoluble

Table 5.5.2 - Ranked second pass screen hits in descending order of priority. Ranking was arbitrarily done based on the respective mean ANI increase and spread (as indicated by the standard error of the mean (SEM)) of the processing competent and deficient constructs. For visualisation, these data are amongst those plotted in panel B of figure 5.5.1, which also includes statistical analyses. Hit peptides and scrambles thereof corresponding to different regions of the INbox were synthesized with a view to carrying out further characterisation.

5.6 DISCUSSION

The INbox library screen had two goals: firstly, to enquire whether or not the Aurora B-INbox interaction can be easily modulated and, secondly, it was intended to serve as a proof-of-concept for the use of SICLOPPS-derived libraries in human cells. With regards to the first goal, it is not possible to reach conclusions about the amenability of the Aurora B-INbox interaction to inhibition with small drug-like compounds at this juncture. Although a number of statistically significant and reproducible hits were obtained following two rounds of screening, their effects were disappointingly modest. The top ranked library screen hit elicited an ANI increase of 1.7X fold relative to the empty vector control in the second pass validation screen. In contrast to this, the SICLOPPS^{T69AH72A} version of the whole INbox construct (i.e. that with the greatest activity) elicited a 7.9X increase. Although it is expected that the unprocessed whole INbox construct will be present in higher concentrations in cells than library-derived CPs, the low relative activity of library constructs remains disappointing. Furthermore, limitations of the SICLOPPS approach have impeded the comparison of hit peptide activities and the assaying of increasing CP concentrations. Given the preliminary nature of these results, firm conclusions regarding whether the Aurora B-INbox interaction can be modulated with small drug-like compounds cannot be drawn from the evidence here.

In discussing the experiments contained within this chapter from the perspective of a proof-of-concept study, I will begin with reflections on target selection, library design and screen strategy. The latter section will focus a reassessment of the initial screen protocol in light of subsequent improvements. This will be followed by a discussion of the results of both phases of the screen. Conclusions on the use of SICLOPPS technology, of

combinatorial versus targeted library design approaches, and of screening in human cells will be made in the next chapter.

TARGET SELECTION

The decision to target the Aurora B-INbox interaction was made early on with a view to evaluate the potential of screening genetically-encoded cyclic peptide libraries in human cells. Having been the subject of intensive research, structural interactions between CPC were already well characterized. Furthermore, widespread cytokinesis failure is a hallmark of CPC loss of function that is readily detectable by microscopy or flow cytometry; this readout was thus amenable to cell-based screens. This made the targeting of interactions within CPC attractive candidates.

Although a number of survivin binders have been developed, targeting the Aurora B-INbox interaction would constitute a novel means of quelling CPC activity. Although it had been suggested that this interaction could be a therapeutically important target (Keen and Taylor, 2009), no agents targeting it have been reported in the literature. The subsequent experiments in which I expressed the whole INbox domain in HeLa cells (presented in Chapter 4) have supported the targeting of this interaction as a means of abrogating CPC activity.

Lastly, the size range of CPs capable of being translocated across the nuclear membrane and/or the nuclear pore complex is currently uncertain. As nuclear envelope breakdown occurs early in mitosis, the selection of a mitotic target circumvented this potential barrier and was therefore attractive for a proof-of-concept study.

DESIGN

As we opted to base the library design on primary amino acid sequence homology (discussed further below), continuous epitopes were desirable. Thanks to its extended conformation, the INbox presented a far more attractive candidate surface than the complimentary surface on Aurora B. In the latter instance, many topologically adjacent residues are not contiguous.

The region of the INbox selected when designing the library was based on the fragment used by Sessa and colleagues in their 2005 crystallography study. This study reported the structure of *Xenopus laevis* Aurora B and INbox in complex with the ATP analogue inhibitor hesperidin and was the only such study published at that time. Subsequently to the design and generation of the library, a crystal structure of the human Aurora B-INbox with VX-680 was published (Elkins et al., 2012). Thankfully, there is good concordance between the two studies regarding the span of the INbox stably associated with the Aurora B N-terminal lobe.

Two independent prior studies had demonstrated the importance of residues homologous to W845 and F881 for Aurora B activation (Sessa et al., 2005; Xu et al., 2009). That INCENP mutants lacking these residues cannot rescue Aurora B activation highlighted these regions as sites of interest. Sub-libraries were thus created centered on residues W845 and F881, with an enhanced focus on W845 as the function of INCENP^{W845G} was known to be more severely compromised than INCENP^{F881A}.

The decision to focus on these residues and not, for example, the TSS motif was supported by the whole INbox mutagenesis study I subsequently carried out (presented in chapter 4). However, the decision to create an additional pentapeptide library centered on W845 but not F881A was not supported by the results of these experiments as both mutants lost their dominant negative effect to the same extent.

Given the modest activities detected in assays so far, it appears that directly templating the library design directly on primary amino acid sequence may have been overly simplistic. Notably, screens of fully randomised cyclic peptide or aptamer libraries against two hybrid-based reporter often yield hits with little sequence homology to either bait protein. Beyond the sequence of the variable region within the library, it is also important to consider other features that may have improved activity or comparability: namely, peptide size and the possible addition of fixed accessory residues to promote processing.

Designing a library of CP-generating constructs rather than aptamer-like SICLOPPS^{T69AH72A} versions thereof (i.e. presenting the construct as a loop flanked by two intein moieties) may seem ill advised in light of results of the whole INbox expression. In those experiments, the ANI increase was greatest in the SICLOPPS^{T69AH72A} versions of INbox^{WT} and INbox^{TSS}. As I have argued, in that instance, the dominant negative effects observed were likely to have been elicited by the linear precursors, not the CP products. There are, however, numerous reasons why the INbox CPs may not have had the anticipated effect (discussed in section 4.5). One of the aims of the library screen had been to determine whether the Aurora B-INbox interaction can be modulated with compounds with drug-like properties. For this reason,

expressing peptide epitopes attached to a large intein moiety was not desirable. Thus, the size of the CPs screened was kept small.

When designing the W845 and F881 sub-libraries, the intention was to enhance chances of producing an effective peptide by varying conformational constraint and promoting processing of peptides associated with key epitopes. The first would be achieved by having related peptide sequences present on varying cyclic peptide backbone lengths. By varying cyclic peptide lengths, peptides could be generated containing similar sequences but with different residues in extein positions known to modulate processing efficiencies. Of the 10 library constructs eliciting the greatest ANI increase in the first pass screen, 8 did not lie within the 7 variable residue core library. This supports the decision to make sub libraries with varying cyclic peptide length.

An alternative design strategy to promote CP processing would have been to flank library sequences with extein residues reported to promote processing (Scott et al., 2001). That is, further to the invariable cysteine residue in position +1 of library CPs, additional fixed non-INbox residues would be added to the start and end of each CP. This strategy was considered but not pursued at the time of library design. The main reasons for this were, firstly, that determining the best combination of residues would have been time consuming and, secondly, that this would have added unwanted bulk to the peptides. However, processing controls carried out during the second phase of the screen highlighted unexpectedly high processing heterogeneity within the library. In hindsight, this could have been a good strategy to promote both overall processing and more homogeneous, thus more comparable, CP production levels. Another strategy that could also be

combined with this one is the use of other inteins reported to be more efficient. This is explored in section 6.3 of the next chapter.

CONSTRUCTION

Unlike combinatorial libraries, the cost of having even a relatively small library of discrete constructs generated commercially was disappointingly prohibitive. However, the process of manually constructing a library from long oligos proved straightforward and was completed in a matter of weeks. Unlike generating combinatorial libraries whose complexity and size can be dramatically curtailed by any inefficiency in the cloning process (R. Almeida, personal communications), there were no requirements for extensive optimization in this instance.

In constructing the library, only three individual clones could not be obtained despite several attempts to do so. Unfortunately, all three overlap with one or the other conserved residue targeted in the design of the library (i.e. W845 and F881). Failure to obtain these clones despite good yields of PCR product could be due to a number of reasons including cytotoxicity in *E. coli*. However, the purpose of the initial library screen was to identify hits, not to compare the respective activities a continuous series of INbox fragments. As such, the failure to obtain a small number of the desired clones was not deemed problematic. Despite apparent questions as to their suitability, the manual construction of discrete libraries on this scale is simple and not a barrier in itself.

LIBRARY SCREENING APPROACH AND RESULTS

The smaller size and discrete nature of the rationally designed library was attractive as it allowed for a greater range of screen readouts (i.e. these were not limited to strong positive selection) and made screening in human

cells feasible. The rationale, advantages and drawbacks of read out selection, carrying out a microscopy – rather than a flow cytometry – based screen, and screening by transient transfection have already been discussed in Chapter 3. Broader discussions regarding phenotypic screening of libraries for PPI modulators in human cells will be presented in the Conclusions chapter. Here I will focus first on the suitability of the various iterations of the library screen protocol, followed by the limitations imposed by the SICLOPPS technology.

As mentioned, the first-pass screen was carried out early in this project and with a library of 97 members. Due to the high number of constructs to be screened, the initial screening strategy was designed as to minimize time and labour required. Lessons learnt from this initial screen led to improvements in the experimental and scoring protocols. The revised protocol was employed in all other ANI assays present within this work including the finer-grained second pass screen.

Two key improvements that were introduced as the assay was being developed were the use of phalloidin staining and blind scoring. The first served to improve scoring accuracy, the latter scoring reliability. Come the second phase of the screen, experiments were carried out using a clonal HeLa Kyoto isolate (refer back to section 3.3) with a lower background ANI than the original cell line. Important changes were also made to the scoring methodology: constructs were assayed in triplicate to permit statistical analyses, and >1000 cells were scored instead of 5 fields of view (the rationale for the former is presented in section 3.3).

A shortcoming of scoring fixed numbers of fields of view is that this leads to the systematic scoring of fewer cells in conditions that negatively affect cell division rates and/or cell survival. This is a valid concern as the number of cells scored in the initial screen ranged from 135 to 580 per

replicate (average=344.4, SD=85.6). Furthermore, second pass screen samples were assayed in triplicate, thereby permitting statistical analyses.

When comparing the same constructs, the rate of reproducibility observed with the more stringent protocol was 70%. With the caveat that the false negative and positive rates in the final protocol are unknown, this suggests a moderate false positive rate in the initial screen. In hindsight, the initial protocol may have been insufficiently stringent, even for a rough initial screen.

Limitations of assaying genetically encoded constructs through transient transfections were first discussed in section 4.5. Although the effects of the INbox^{WT} and INbox^{AAA} constructs were readily apparent after 48 hours (recall Chapter 4), hit peptides had very modest effects over the same time period. Although the effects of hit peptide ANI increases were statistically significant under the screen conditions, it would have been informative to also assay their effect over longer time periods. However, the expression of transiently transfected SICLOPPS constructs might not persist sufficiently long to permit this.

A possible explanation for the low potency of library constructs would be that the INbox fragments assayed in the library were simply too short to effectively recapitulate the interactions made by INbox^{WT} and INbox^{AAA}. Although processing had a marked negative impact on the potency of the INbox^{WT} and INbox^{AAA} constructs this effect was not clearly apparent during the validation of library constructs. The ANI increases elicited by most of the validated hits were indistinguishable in their processing competent and deficient forms.

Comparisons of hit activity proved extremely difficult using SICLOPPS constructs. The greatest complication was the range of processing efficiency within the library, which exceeded my expectations. However, the lack of correlation between processing and ANI increase suggests that the effect elicited by hits is not due to circularization efficiency alone. It is therefore likely that the ANI increases detected are sequence specific. The processing heterogeneity hampered the comparison of the hits obtained in this screen but also dissuaded testing whether the effect of hit constructs is sequence dependent using SICLOPPS. This is because scrambled or mutant (e.g. creating the mutant equivalent to W845G in constructs with sequence homology to this region of the INbox) versions of hit constructs might not be processed to the same extent as the original, which would prevent direct comparison. Alternative means of comparing hit peptides directly against modified versions thereof are necessary (see Perspectives chapter).

The phenotypic effect elicited by any true hit should also correlate with CP levels. Processing heterogeneity and technical constraints prohibiting direct CP detection create significant hurdles to testing the correlation between phenotype severity and CP levels using SICLOPPS methodology. However, that between phenotype and SICLOPPS global expression (i.e. not taking into account processing) levels could be verified thanks to the epitope tag present on the SICLOPPS backbone. Another means of testing this is by assaying different concentrations of the peptides. Although achievable through inducible expression systems, this would still be complicated by processing heterogeneity. For this reason, we are currently pursuing the microinjection of synthetic versions of hit peptides and scrambled versions thereof, which would serve a negative controls (see Perspectives chapter).

The modest activity of hit constructs makes it difficult to verify their effect on the CPC. This is another area of interest that could be investigated through the use of chemically synthesized CPs. Nevertheless, some of the results obtained suggest that the hits identified might be specific to the CPC. Firstly, CP activity appears to be sequence dependent, as evidenced by the lack of correlation between processing efficiency and activity. Secondly, hit peptides sequences clustered into two regions, each encompassing INbox W845 or F881 – two INbox residues that are each required for INCENP binding to Aurora B (Section 4.3). Although conclusions regarding the amenability of the Aurora B-INbox interaction as a drug target cannot be made at this stage, these are small but encouraging signs.

CHAPTER 6: CONCLUSIONS

6.1 TARGETTING THE INBOX/AURORA B INTERACTION

CPC function is essential for cell division and survival. Numerous prior studies had demonstrated that the presence and function of all CPC components are essential. Although loss of the CPC at centromeres can be tolerated in tissue culture, failure to achieve CPC localization to the anaphase spindle and midbody correlates with cytokinesis failure and loss of viability. The evidence generated in the present study further supports the hypothesis that directly blocking the interaction between Aurora B and INCENP alone is sufficient to impair CPC function and the completion of cell division. Expression of the linear full-length INbox resulted in a highly significant increase in multinucleate and micronucleate cells, which is consistent with CPC loss-of-function.

Failure to correctly complete cell division occurred despite the presumed maintenance of both endogenous Aurora B and INCENP protein levels. The respective localization of INCENP, survivin, and borealin, and the activation state of Aurora B, under these conditions remain to be verified. Such observations might potentially yield new insights into the biological role and function of the CPC.

In partial agreement with previously published experimental results, the failure of two point mutants of the INbox – W845G and F881A – to affect CPC activity when overexpressed in isolation or combination indicates that both sites are essential for INbox binding to Aurora B. The sum of these and previous results suggests that blocking either of these contact points between Aurora B and INCENP might be sufficient to impair CPC function. The structure of the INbox is extended but the complimentary interaction

surface on the Aurora B N-terminal lobe consists of a shallow groove that contains W845 and F881 interaction pockets.

Attempts to determine whether the Aurora B-INbox interaction is one that can be modulated with compounds smaller than the whole INbox were inconclusive (see below). A small number of very modest but statistically significant leads were obtained but their activity remains to be verified (see 'Perspectives' section). Thus, it is not possible to draw conclusions on whether this interaction is amenable to being drugged at this juncture.

Should the interaction prove amenable to inhibition with small drug-like compounds, this would open up the possibility of developing inhibitors that might offer new research and therapeutic avenues. Although potentially novel insights remain to be gained from the genetically encoded whole INbox constructs, drugs allow for a far superior degree of temporal and concentration control that might open further experimental avenues and might transfer to the clinic.

6.2 RATIONALLY-DESIGNED CP LIBRARIES & SCREENING IN HUMAN CELLS

From the perspective of carrying out a feasibility study, several features of using the Aurora B-INbox interaction presented challenges. Uncertainty about the effects of directly blocking this interaction, and their scale, in cell culture made it difficult to optimize the screen and assess its results. When the CPC function assay and the screen were devised, no reagents capable of directly inhibiting the endogenous Aurora B-INbox interaction were available. The dominant negative effect of the control used in lieu of this – overexpression of a dominant-negative Aurora B mutant – was highly pronounced and expected to be substantially more severe than

that of dissociating the Aurora B-INbox interaction (as discussed in section 3.5 and 4.5). The strongest hits obtained after the initial phase of the INbox library screen were very modest in comparison to this negative control, but this was expected. However, the full-length INbox overexpression experiments (Chapter 4) had not yet been performed at that time. It was therefore not possible to gauge the relevance of hit activities compared to this more representative control prior to proceeding with the second phase of the screen.

Another challenge of using the Aurora B-INbox for this pilot study was the negative impact of its inhibition on cell survival. This effect was strongly anticipated at the time the screen was devised and subsequently confirmed by the experiments in which I overexpressed the whole INbox (Chapter 4). When screening genetically encoded libraries, readouts inconsistent with cell survival complicate the recovery of DNA sequences associated with the desired phenotype as cells expressing the strongest hits are missing from the read-out. Simply transposing and testing established SICLOPPS library design in human cells alone would not have been sufficient – new approaches had to be devised.

Whilst a more in depth assessments of hit peptides (see Perspectives) is needed in order to verify their observed activity, it was clear that their effect in the cell-based assay was far less pronounced than that of the whole INbox peptide. Designing a small library of discrete members presented numerous practical advantages over a pooled library. Amongst them, library construction and tissue culture and transfection conditions were relative easy to optimize. Furthermore, the discrete nature of the library and its small size allowed for each construct to be assayed in isolation. This circumvented the need to recover hit peptide sequences and permitted the design of an *in vivo*

screen against an essential interaction. However, the costs of trading off chemical diversity for practicality may have been too high. The disappointingly modest leads obtained in this study raise serious questions about the suitability of the small, targeted, SICLOPPS library approach employed in this study.

Acknowledging the shortcomings of basing a feasibility study on the use of SICLOPPS technology in human cells on a then untested interaction, I must nevertheless conclude that there remain significant hurdles to the effective screening SICLOPPS libraries in this system. I must advise that small rationally designed libraries based on the approach used in this study should be discouraged. Iteratively assaying much larger INbox fragments in their linear and circular forms in combination with flexible linkers might have yielded a shortened construct capable of reiterating the dominant negative effect of the larger INbox construct. If restricting the size of the cyclic peptides is desired, a large combinatorial library (i.e. one in which all or some of the peptide residues are randomised) might be better suited. Cell-based screening of such large libraries remains best suited to lower eukaryotes and prokaryotes as these are more easily grown and manipulated.

With regards to identifying small CPs capable of inhibiting the Aurora B-INbox interaction with sizes that would indicate that this interaction could be inhibited with drugs, the alternative combinatorial approaches described above would have been advantageous. Prokaryotes represent a heterologous system if screening for eukaryotic PPI inhibitors. However, the ease and speed with which genetically encoded libraries can be screened in prokaryotic systems outweigh this concern. A possible discovery pipeline might first involve prokaryotic screens as a rapid and high throughput means of the identifying, optimizing and prioritizing a panel of strong initial hits.

These hits could then be subsequently assayed in the endogenous context through more suitable means, such as by using chemically synthesized CPs. This approach is exemplified by the 2008 study by Tavassoli and colleagues in which cyclic peptide inhibitors of the interaction between HIV Gag protein and human TSG101 (tumour susceptibility gene 101) were first identified using a bacterial reverse-two hybrid reporter (Tavassoli et al., 2008). Cell-penetrant versions of hit peptides were then chemically synthesized and assayed against a reporter of HIV budding in 293T cells.

The use of bacterial reverse-two hybrid reporters might have been particularly well suited to finding small peptide inhibitors of the Aurora B-INbox interaction. Given the absence of the CPC in prokaryotes, the interaction can be reiterated within a reverse two-hybrid reporter system generating strong positive selection for hits. An additional advantage of reverse-two hybrid reporters is that, unlike a phenotypic screen, these are inherently targeted. As discussed previously (section 3.5), failure to elicit an ANI increase in the CPC function assay is consistent with intact CPC function but the converse is not necessarily true. This is because many other factors than the CPC are involved in chromosome segregation and cytokinesis.

6.3 REASSESSING SICLOPPS TECHNOLOGY

Many features of the SICLOPPS technology make this approach very attractive for identifying CP modulators using combinatorial libraries. However, several drawback and limitations of the technology became apparent during this present study. The use of genetically encoded constructs with varying rates of processing ultimately frustrated efforts to compare hit peptide activities side-by-side, and rank these. As already mentioned, SICLOPPS methodology would therefore best be deployed in the early stages

of a discovery project with plans already in place to generate hit peptides through other means for subsequent assays.

The observed variation in processing efficiency means that, in terms of CP production, the library screened in this study may have been effectively smaller than that designed. More generally, this raises concerns that the effective chemical diversity contained within SICLOPPS-derived combinatorial libraries may be overestimated and overstated.

It is noteworthy that all studies reporting the use of split inteins to generate CPs *in vivo* have thus far relied on inteins that were amongst the first described. The *Ssp DnaE* intein-mediated protein *trans* splicing reaction has been reported to have a half-life of 175 minutes at 23^{o}C (Martin et al., 2001). As already discussed, the processing of the *Ssp DnaE* intein can also be very sensitive to extein residues. Since then, inteins reported to possess much faster processing speeds and greater tolerance for variation in extein residues have been described. Of particular note is the *dnaE* intein from *Nostoc Punctiforme* (*Npu*), which is reported to undergo *trans* splicing with a half life of 63 seconds (Zettler et al., 2009). The study presented here was done using the *Ssp DnaE* intein as it was available from our collaborator and is still commonly used to generate cyclic peptide libraries. Newly characterised inteins, however, have the potential for improving both SICLOPPS processing speeds and extent, and may yield better CP libraries.

CHAPTER 7: PERSPECTIVES

The results presented above are but parts of stories that can be further developed. As discussed in the previous chapter, two avenues that could warrant further exploration are the biological consequences of full length INbox expression and further validation of the effects of hit peptides identified in the INbox screen.

Overexpression of full length INbox caused a significant increase in cytokinesis failure and alignment defects. These effects are likely to be inconsistent with cell survival over longer periods of time. Our results support the targeting of the Aurora B-INbox interaction as a means of abrogating Aurora B activity. They also suggest that the regions interacting with not only hINCENP W845 but also F881 could be suitable target sites for drug discovery efforts.

It is interesting that, by using point mutants of the dominant negative INbox construct, we implicated hINCENP F881 in Aurora B binding *in vivo*. This is in contrast to a previous study using heterologously expressed human proteins that found this residue to be required for activation, but not binding of Aurora B (Honda et al., 2003). The difference in the results obtained using our approach – i.e. abolishing a dominant negative effect – and *in vitro* assays suggests that these might differ in their sensitivity.

At the present juncture, the INbox SICLOPPS library screen has yielded a series of hits with only moderate activity in the CPC function assay. To further investigate these, three of these hits have been chemically synthesized along with matched scrambled sequence peptides by our collaborator Daniel St Cyr at the Université de Montréal (recall Table 5.5.2). The use of chemically-synthesized peptides will permit the control of

experimental concentrations, something that proved elusive using SICLOPPS. However, the very modest effect of the hit peptides identified should be reiterated – these are probably not be worth pursuing outside the scope of this study.

Before commenting on the wider significance of our findings concerning the targeting of the Aurora B-INbox interaction, is worth remembering that all experiments in study were carried out using HeLa Kyoto or a clonal isolate thereof. Although deemed suitable for this study (discussed in section 3.2), HeLa cells are aberrant. Recent genomic and transcriptomic analysis of the cell line has revealed massive genomic rearrangements and differential expression of several essential pathways (Landry et al., 2013). The generalisability of our findings should therefore be verified in a range of cell lines, including non-transformed cells.

Overall, our results support the targeting of the Aurora B-INbox interaction but are inconclusive regarding its amenability to inhibition with small molecules or peptides. The potential novel mechanism for Aurora B inhibition is highly interesting in itself. In light of the debate surrounding the effectiveness of targeting mitosis in cancer therapy that has since emerged (Komlodi-Pasztor et al., 2011; Mitchison, 2012), the therapeutic potential of this approach appears to be less promising than when this study began. Inhibitors of the Aurora B-INbox interaction could nonetheless be useful for research applications, such as in dissecting any possible CPC-independent role of Aurora B.

REFERENCES

Abel-Santos, E., Scott, C.P., and Benkovic, S.J. (2003). Use of inteins for the in vivo production of stable cyclic peptide libraries in E. coli. Methods Mol Biol *205*, 281-294.

Abramoff, M.D., Magalhães, P.J., Ram, S.J. (2004). Image processing with ImageJ, pp. 36--42.

Adams, R.R., Maiato, H., Earnshaw, W.C., and Carmena, M. (2001). Essential roles of Drosophila inner centromere protein (INCENP) and aurora B in histone H3 phosphorylation, metaphase chromosome alignment, kinetochore disjunction, and chromosome segregation. The Journal of cell biology *153*, 865-880.

Adams, R.R., Wheatley, S.P., Gouldsworthy, A.M., Kandels-Lewis, S.E., Carmena, M., Smythe, C., Gerloff, D.L., Earnshaw, W.C. (2000). INCENP binds the Aurora-related kinase AIRK2 and is required to target it to chromosomes, the central spindle and cleavage furrow. Current biology : CB *10*, 1075--1078.

Ainsztein, A.M., Kandels-Lewis, S.E., Mackay, A.M., Earshawn, W.C. (1998). INCENP Centromere and Spindle Targeting: Identification of Essential Conserved Motifs and Involvement of Heterochromatin Protein HP1. The Journal of Cell Biology *143*, 1763--1774.

Ambrosini, G., Adida, C., Altieri, D.C. (1997). A novel anti-apoptosis gene, survivin, expressed in cancer and lymphoma. Nature Medicine *3*, 917--921.

Aramburu, J., Yaffe, M.B., López-Rodríguez, C., Cantley, L.C., Hogan, P.G., and Rao, A. (1999). Affinity-driven peptide selection of an NFAT inhibitor more selective than cyclosporin A. Science (New York, NY) *285*, 2129-2133.

Avo Santos, M., van de Werken, C., de Vries, M., Jahr, H., Vromans, M.J.M., Laven, J.S.E., Fauser, B.C., Kops, G.J., Lens, S.M., and Baart, E.B. (2011). A role for Aurora C in the chromosomal passenger complex during human preimplantation embryo development. Human reproduction (Oxford, England) *26*, 1868-1881.

Bain, J., Plater, L., Elliott, M., Shpiro, N., Hastie, C.J., McLauchlan, H., Klevernic, I., Arthur, J.S., Alessi, D.R., Cohen, P. (2007). The selectivity of protein kinase inhibitors: a further update. The Biochemical journal *408*, 297--315.

Balboula, A.Z., Schindler, K. (2014). Selective disruption of aurora C kinase reveals distinct functions from aurora B kinase during meiosis in mouse oocytes. PLoS genetics *10*, e1004194.

Banks, D.P., Plescia, J., Altieri, D.C., Chen, J., Rosenberg, S.H., Zhang, H., Ng, S.C. (2000). Survivin does not inhibit caspase-3 activity. Blood *96*, 4002--4003.

Becker, M., Stolz, A., Ertych, N., Bastians, H. (2010). Centromere localization of INCENP-Aurora B is sufficient to support spindle checkpoint function. Cell cycle (Georgetown, Tex) *9*, 1360--1372.

Bishop, J.D., and Schumacher, J.M. (2002). Phosphorylation of the carboxyl terminus of inner centromere protein (INCENP) by the Aurora B Kinase stimulates Aurora B kinase activity. The Journal of biological chemistry *277*, 27577-27580.

Boice, J.D., Day, N.E., Andersen, A., Brinton, L.A., Brown, R., Choi, N.W., Clarke, E.A., Coleman, M.P., Curtis, R.E., Flannery, J.T., et al. (1985). Second cancers following radiation treatment for cervical cancer. An international collaboration among cancer registries. Journal of the National Cancer Institute *74*, 955--975.

Bolton, M.A., Lan, W., Powers, S.E., McCleland, M.L., Kuang, J., and Stukenberg, P.T. (2002). Aurora B kinase exists in a complex with survivin and INCENP and its kinase activity is stimulated by survivin binding and phosphorylation. Molecular biology of the cell *13*, 3064-3077.

Bourhis, E.a.H.S.G.a.C.A.G. (2007). The mitotic regulator Survivin binds as a monomer to its functional interactor Borealin. The Journal of biological chemistry *282*, 35018--35023.

Bourhis, E., Hymowitz, S.G., Cochran, A.G. (2009). Phosphorylation of a borealin dimerization domain is required for proper chromosome segregation. Biochemistry *48*, 6783--6793.

Brenner, D.J., Curtis, R.E., Hall, E.J., Ron, E. (2000). Second malignancies in prostate carcinoma patients after radiotherapy compared with surgery. Cancer *88*, 398--406.

Butcher, E.C. (2005). Can cell systems biology rescue drug discovery? Nature reviews Drug discovery *4*, 461-467.

Capalbo, L., Montembault, E., Takeda, T., Bassi, Zi., Glover, D.M., D'avino, P.P. (2012). The chromosomal passenger complex controls the function of

endosomal sorting complex required for transport-III Snf7 proteins during cytokinesis. Open Biology *2*, 120070--120070.

Cardozo, A.K., Buchillier, V., Mathieu, M., Chen, J., Ortis, F., Ladrière, L., Allaman-Pillet, N., Poirot, O., Kellenberger, S., Beckmann, J.S.*, et al.* (2007). Cell-permeable peptides induce dose- and length-dependent cytotoxic effects. Biochimica et biophysica acta *1768*, 2222-2234.

Carlton, J.G., Caballe, A., Agromayor, M., Kloc, M., Martin-Serrano, J. (2012). ESCRT-III governs the Aurora B-mediated abscission checkpoint through CHMP4C. Science (New York, NY) *336*, 220--225.

Carmena, M., Ruchaud, S., and Earnshaw, W.C. (2009). Making the Auroras glow: regulation of Aurora A and B kinase function by interacting proteins. Current opinion in cell biology *21*, 796-805.

Carmena, M., Wheelock, M., Funabiki, H., Earnshaw, W.C. (2012). The chromosomal passenger complex (CPC): from easy rider to the godfather of mitosis. Nature reviews Molecular cell biology *13*, 789--803.

Carvalho, A., Carmena, M., Sambade, C., Earnshaw, W.C., Wheatley, S.P. (2003). Survivin is required for stable checkpoint activation in taxol-treated HeLa cells. Journal of cell science *116*, 2987--2998.

Chantalat, L., Skoufias, D.A., Kleman, J.P., Jung, B., Dideberg, O., Margolis, R.L. (2000). Crystal structure of human survivin reveals a bow tie-shaped dimer with two unusual alpha-helical extensions. Molecular cell *6*, 183--189.

Chen, H.-L., Tang, C.-J.C., Chen, C.-Y., and Tang, T.K. (2005). Overexpression of an Aurora-C kinase-deficient mutant disrupts the Aurora-B/INCENP complex and induces polyploidy. Journal of biomedical science *12*, 297-310.

Chen, L., Zhang, Y., Li, G., Huang, H., and Zhou, N. (2010). Functional characterization of a naturally occurring trans-splicing intein from Synechococcus elongatus in a mammalian cell system. Analytical biochemistry *407*, 180-187.

Chen, Q., Lakshmikanth, G.S., Spudich, J.A., De Lozanne, A. (2007). The localization of inner centromeric protein (INCENP) at the cleavage furrow is dependent on Kif12 and involves interactions of the N terminus of INCENP with the actin cytoskeleton. Molecular biology of the cell *18*, 3366--3374.

Cheng, L., Naumann, T.A., Horswill, A.R., Hong, S.-J., Venters, B.J., Tomsho, J.W., Benkovic, S.J., and Keiler, K.C. (2007). Discovery of antibacterial cyclic

peptides that inhibit the ClpXP protease. Protein science : a publication of the Protein Society *16*, 1535-1542.

Clamp, M., Cuff, J., Searle, S.M., and Barton, G.J. (2004). The Jalview Java alignment editor. Bioinformatics *20*, 426-427.

Cochran, A. (2000). Antagonists of protein-protein interactions. Chemistry & Biology *7*, R85-R94.

Colas, P., Cohen, B., Jessen, T., Grishina, I., McCoy, J., and Brent, R. (1996). Genetic selection of peptide aptamers that recognize and inhibit cyclin-dependent kinase 2. Nature *380*, 548-550.

Cooke, C.A., Heck, M.M., and Earnshaw, W.C. (1987). The inner centromere protein (INCENP) antigens: movement from inner centromere to midbody during mitosis. The Journal of cell biology *105*, 2053-2067.

Gene Codes Corporation, G.C. Sequencher® version 5.0 sequence analysis software (Ann Arbor, MI USA).

Craik, D.J. (2006). Chemistry. Seamless proteins tie up their loose ends. Science (New York, NY) *311*, 1563-1564.

Dai, J., Sultan, S., Taylor, S.S., Higgins, J.M.(2005). The kinase haspin is required for mitotic histone H3 Thr 3 phosphorylation and normal metaphase chromosome alignment. Genes & development *19*, 472--488.

Ditchfield, C., Johnson, V.L., Tighe, A., Ellston, R., Haworth, C., Johnson, T., Mortlock, A., Keen, N., and Taylor, S.S. (2003). Aurora B couples chromosome alignment with anaphase by targeting BubR1, Mad2, and Cenp-E to kinetochores. The Journal of cell biology *161*, 267-280.

Douglas, M.E., Davies, T., Joseph, N., Mishima, M. (2010). Aurora B and 14-3-3 coordinately regulate clustering of centralspindlin during cytokinesis. Current biology : CB *20*, 927--933.

Drechsel, D.N., Hyman, A.A., Hall, A., Glotzer, M. (1997). A requirement for Rho and Cdc42 during cytokinesis in Xenopus embryos. Current biology : CB *7*, 12--23.

Du, J., Kelly, A.E., Funabiki, H., Patel, D.J. (2012). Structural basis for recognition of H3T3ph and Smac/DIABLO N-terminal peptides by human Survivin. Structure (London, England : 1993) *20*, 185--195.

Elkins, J.M., Santaguida, S., Musacchio, A., Knapp, S. (2012). Crystal structure of human aurora B in complex with INCENP and VX-680. Journal of medicinal chemistry *55*, 7841--7848.

Emanuele, M.J., Lan, W., Jwa, M., Miller, S.A., Chan, C.S., Stukenber, P.T. (2008). Aurora B kinase and protein phosphatase 1 have opposing roles in modulating kinetochore assembly. The Journal of cell biology *181*, 241--254.

Failes, T.W., Mitic, G., Abdel-Halim, H., Po'uha, S.T., Liu, M., Hibbs, D.E., and Kavallaris, M. (2012). Evolution of resistance to aurora kinase B inhibitors in leukaemia cells. PloS one *7*, e30734-e30734.

Fei, F., Stoddart, S., Groffen, J., Heisterkamp, N. (2010). Activity of the Aurora kinase inhibitor VX-680 against Bcr/Abl-positive acute lymphoblastic leukemias. Molecular cancer therapeutics *9*, 1318--1327.

Fernández-Miranda, G., Trakala, M., Martín, J., Escobar, B., González, A., Ghyselinck, N.B., Ortega, S., Cañamero, M., Pérez de Castro, I., and Malumbres, M. (2011). Genetic disruption of aurora B uncovers an essential role for aurora C during early mammalian development. Development (Cambridge, England) *138*, 2661-2672.

Frankel, A.D., and Pabo, C.O. (1988). Cellular uptake of the tat protein from human immunodeficiency virus. Cell *55*, 1189-1193.

Gassmann, R., Carvalho, A., Henzing, A.J., Ruchaud, S., Hudson, D.F., Honda, R., Nigg, E.A., Gerloff, D.L., and Earnshaw, W.C. (2004). Borealin: a novel chromosomal passenger required for stability of the bipolar mitotic spindle. The Journal of cell biology *166*, 179-191.

Giet, R., and Glover, D.M. (2001). Drosophila aurora B kinase is required for histone H3 phosphorylation and condensin recruitment during chromosome condensation and to organize the central spindle during cytokinesis. The Journal of cell biology *152*, 669-682.

Girdler, F., Gascoigne, K.E., Eyers, P.A., Hartmuth, S., Crafter, C., Foote, K.M., Keen, N.J., and Taylor, S.S. (2006). Validating Aurora B as an anti-cancer drug target. Journal of cell science *119*, 3664-3675.

Girdler, F., Sessa, F., Patercoli, S., Villa, F., Musacchio, A., and Taylor, S. (2008). Molecular basis of drug resistance in aurora kinases. Chemistry & biology *15*, 552-562.

Glover, D.M., Leibowitz, M.H., McLean, D.A., Parry, H. (1995). Mutations in aurora prevent centrosome separation leading to the formation of monopolar spindles. Cell *81*, 95--105.

Gottesman, M.M. (2002). Mechanisms of cancer drug resistance. Annual review of medicine *53*, 615--627.

Hagting, A., Den Elzen, N., Vodermaier, H.C., Waizenegger, I.C., Peters, J.M., Pines, J. (2002). Human securin proteolysis is controlled by the spindle checkpoint and reveals when the APC/C switches from activation by Cdc20 to Cdh1. The Journal of cell biology *157*, 1125--1137.

Hans, F., Skoufias, D.A., Dimitrov, S., and Margolis, R.L. (2009). Molecular distinctions between Aurora A and B: a single residue change transforms Aurora A into correctly localized and functional Aurora B. Molecular biology of the cell *20*, 3491-3502.

Hardwicke, M.A., Oleykowski, C.A., Plant, R., Wang, J., Liao, Q., Moss, K., Newlander, K., Adams, J.L., Dhanak, D., Yang, J.*, et al.* (2009). GSK1070916, a potent Aurora B/C kinase inhibitor with broad antitumor activity in tissue culture cells and human tumor xenograft models. Molecular cancer therapeutics *8*, 1808-1817.

Hauf, S., Cole, R.W., LaTerra, S., Zimmer, C., Schnapp, G., Walter, R., Heckel, A., van Meel, J., Rieder, C.L., and Peters, J.-M. (2003). The small molecule Hesperadin reveals a role for Aurora B in correcting kinetochore-microtubule attachment and in maintaining the spindle assembly checkpoint. The Journal of cell biology *161*, 281-294.

Hayashi-Takanaka, Y.a.Y.K.a.N.N.a.K.H. (2009). Visualizing histone modifications in living cells: spatiotemporal dynamics of H3 phosphorylation during interphase. The Journal of cell biology *187*, 781--790.

Hernáez, B., Tarragó, T., Giralt, E., Escribano, J.M., and Alonso, C. (2010). Small peptide inhibitors disrupt a high-affinity interaction between cytoplasmic dynein and a viral cargo protein. Journal of virology *84*, 10792-10801.

Holland, A.J.a.C.D.W. (2009). Boveri revisited: chromosomal instability, aneuploidy and tumorigenesis. Nature reviews Molecular cell biology *10*, 478--487.

Honda, R., Körner, R., and Nigg, E.A. (2003). Exploring the functional interactions between Aurora B, INCENP, and survivin in mitosis. Molecular biology of the cell *14*, 3325-3341.

Hoppe-Seyler, F., Crnkovic-Mertens, I., Tomai, E., and Butz, K. (2004). Peptide aptamers: specific inhibitors of protein function. Current molecular medicine *4*, 529-538.

Horswill, A.R., Savinov, S.N., and Benkovic, S.J. (2004). A systematic method for identifying small-molecule modulators of protein-protein interactions. Proceedings of the National Academy of Sciences of the United States of America *101*, 15591-15596.

Huszar, D., Theoclitou, M.-E., Skolnik, J., and Herbst, R. (2009). Kinesin motor proteins as targets for cancer therapy. Cancer metastasis reviews *28*, 197-208.

Hümmer S., Mayer, T.U. (2009). Cdk1 negatively regulates midzone localization of the mitotic kinesin Mklp2 and the chromosomal passenger complex. Current biology : CB *19*, 607--612.

Jeyaprakash, A.A., Basquin C., Jayachandran, U., Conti, E. (2011). Structural basis for the recognition of phosphorylated histone h3 by the survivin subunit of the chromosomal passenger complex. Structure (London, England : 1993) *19*, 1625--1634.

Jeyaprakash, A.A., Klein, U.R., Lindner, D., Ebert, J., Nigg, E.A.., Conti, E. (2007). Structure of a Survivin-Borealin-INCENP core complex reveals how chromosomal passengers travel together. Cell *131*, 271--285.

Jordan, M.A., and Wilson, L. (2004). Microtubules as a target for anticancer drugs. Nature Reviews Cancer *4*, 253--265.

Järver, P., and Langel, U. (2004). The use of cell-penetrating peptides as a tool for gene regulation. Drug discovery today *9*, 395-402.

Kaestner, P., Stolz, A., and Bastians, H. (2009). Determinants for the efficiency of anticancer drugs targeting either Aurora-A or Aurora-B kinases in human colon carcinoma cells. Molecular cancer therapeutics *8*, 2046-2056.

Kaitna, S., Mendoza, M., Jantsch-Plunger, V., and Glotzer, M. (2000). Incenp and an Aurora-like kinase form a complex essential for chromosome segregation and efficient completion of cytokinesis. Current Biology *10*, 1172-1181.

Kanda, A., Kawai, H., Suto, S., Kitajima, S., Sato, S., Takata, T., and Tatsuka, M. (2005). Aurora-B/AIM-1 kinase activity is involved in Ras-mediated cell transformation. Oncogene *24*, 7266-7272.

Kawashima, S.A., Tsukahara, T., Langegger, M., Hauf, S., Kitajima, T.S., Watanabe, Y. (2007). Shugoshin enables tension-generating attachment of kinetochores by loading Aurora to centromeres. Genes \& development *21*, 420--435.

Kawashima, S.A., Yamagishi, Y., Honda, T., Ishiguro, K., Watanabe, Y. (2010). Phosphorylation of H2A by Bub1 prevents chromosomal instability through localizing shugoshin. Science (New York, NY) *327*, 172--177.

Keen, N., and Taylor, S. (2004). Aurora-kinase inhibitors as anticancer agents. Nature reviews Cancer *4*, 927-936.

Keen, N., and Taylor, S. (2009). Mitotic drivers--inhibitors of the Aurora B Kinase. Cancer metastasis reviews *28*, 185-195.

Kelemen, B.R., Hsiao, K., and Goueli, S.A. (2002). Selective in vivo inhibition of mitogen-activated protein kinase activation using cell-permeable peptides. The Journal of biological chemistry *277*, 8741-8748.

Kelly, A.E., Ghenoiu, C., Xue, J.Z., Zierhut, C., Kimura, H., and Funabiki, H. (2010). Survivin reads phosphorylated histone H3 threonine 3 to activate the mitotic kinase Aurora B. Science (New York, NY) *330*, 235-239.

Kelly, A.E., Sampath, S.C., Maniar, T.A., Woo, E.M., Chait, B.T., Funabiki, H. (2007). Chromosomal enrichment and activation of the aurora B pathway are coupled to spatially regulate spindle assembly. Developmental cell *12*, 31--43.

Khan, A.R., Parrish, J.C., Fraser, M.E., Smith, W.W., Bartlett, P.A., James, M.N. (1998). Lowering the entropic barrier for binding conformationally flexible inhibitors to enzymes. Biochemistry *37*, 16839--16845.

Kimmins, S., Crosio, C., Kotaja, N., Hirayama, J., Monaco, L., Höög, C., van Duin, M., Gossen, J.A., and Sassone-Corsi, P. (2007). Differential functions of the Aurora-B and Aurora-C kinases in mammalian spermatogenesis. Molecular endocrinology (Baltimore, Md) *21*, 726-739.

King, R.W., Peters, J.M., Tugendreich, S., Rolfe, M., Hieter, P., Kirschner, M.W. (1995). A 20S complex containing CDC27 and CDC16 catalyzes the mitosis-specific conjugation of ubiquitin to cyclin B. Cell *81*, 279--288.

Kinsella, T.M., Ohashi, C.T., Harder, A.G., Yam, G.C., Li, W., Peelle, B., Pali, E.S., Bennett, M.K., Molineaux, S.M., Anderson, D.A.*, et al.* (2002). Retrovirally delivered random cyclic Peptide libraries yield inhibitors of interleukin-4 signaling in human B cells. The Journal of biological chemistry *277*, 37512-37518.

Klein, U.R., Nigg, E.A., and Gruneberg, U. (2006). Centromere targeting of the chromosomal passenger complex requires a ternary subcomplex of Borealin, Survivin, and the N-terminal domain of INCENP. Molecular biology of the cell *17*, 2547-2558.

Komlodi-Pasztor, E., Sackett, D., Wilkerson, J., Fojo, T. (2011). Mitosis is not a key target of microtubule agents in patient tumors. Nature reviews Clinical oncology *8*, 244--250.

Kritzer, J.A., Hamamichi, S., McCaffery, J.M., Santagata, S., Naumann, T.A., Caldwell, K.A., Caldwell, G.A., and Lindquist, S. (2009). Rapid selection of cyclic peptides that reduce alpha-synuclein toxicity in yeast and animal models. Nature Chemical Biology *5*, 655-663.

Kufer, T.A., Silljé, H.H.W., Körner, R., Gruss, O.J., Meraldi, P., and Nigg, E.A. (2002). Human TPX2 is required for targeting Aurora-A kinase to the spindle. The Journal of cell biology *158*, 617-623.

Kulukian, A., Han, J.S., Cleveland, D.W. (2009). Unattached kinetochores catalyze production of an anaphase inhibitor that requires a Mad2 template to prime Cdc20 for BubR1 binding. Developmental cell *16*, 105--117.

Lalonde, M.S., Lobritz, M.A., Ratcliff, A., Chamanian, M., Athanassiou, Z., Tyagi, M., Wong, J., Robinson, J.A., Karn, J., Varani, G.*, et al.* (2011). Inhibition of Both HIV-1 Reverse Transcription and Gene Expression by a Cyclic Peptide that Binds the Tat-Transactivating Response Element (TAR) RNA. PLoS pathogens *7*, e1002038-e1002038.

Landry, J.J., Pyl, P.T., Rausch, T., Zichner, T., Tekkedil, M.M., Stüz, A.M., Jauch, A., Aiyar, R.S., Pau, G., Delhomme, N., Gagneur, J., Korbel, J.O., Huber, W., Steinmetz, L.M. (2013). The genomic and transcriptomic landscape of a HeLa cell line. G3 (Bethesda, Md) *3*, 1213--1224.

Lee, J.J., and Swain, S.M. (2006). Peripheral neuropathy induced by microtubule-stabilizing agents. Journal of clinical oncology : official journal of the American Society of Clinical Oncology *24*, 1633--1642.

Lens, S.M.A., Voest, E.E., and Medema, R.H. (2010). Shared and separate functions of polo-like kinases and aurora kinases in cancer. Nature reviews Cancer *10*, 825-841.

Lens, S.M.A.a.R.J.A.a.V.G.a.S.S.W.a.G.G.a.M.R.H. (2006). Uncoupling the central spindle-associated function of the chromosomal passenger complex from its role at centromeres. Molecular biology of the cell *17*, 1897--1909.

Lens, S.M., Rodriguez, J.A., Vader, G., Span, S.W., Giaccone, G., Medema, R.H. (2003). Survivin is required for a sustained spindle checkpoint arrest in response to lack of tension. The EMBO journal *22*, 2934--2947.

Lewis, K.D., Samlowski, W., Ward, J., Catlett, J., Cranmer, L., Kirkwood, J., Lawson, D., Whitman, E., and Gonzalez, R. (2011). A multi-center phase II evaluation of the small molecule survivin suppressor YM155 in patients with unresectable stage III or IV melanoma. Investigational new drugs *29*, 161-166.

Li, C., Wu, Z., Liu, M., Pazgier, M., Lu, W. (2008). Chemically synthesized human survivin does not inhibit caspase-3. Protein science : a publication of the Protein Society *17*, 1624--1629.

Li, X., Sakashita, G., Matsuzaki, H., Sugimoto, K., Kimura, K., Hanaoka, F., Taniguchi, H., Furukawa, K., Urano, T. (2004). Direct association with inner centromere protein (INCENP) activates the novel chromosomal passenger protein, Aurora-C. The Journal of biological chemistry *279*, 47201--47211.

Lindgren, M., Hällbrink, M., Prochiantz, A., and Langel, U. (2000). Cell-penetrating peptides. Trends in pharmacological sciences *21*, 99-103.

Liu, D., Vader, G., Vromans, M.J.M., Lampson, M.A., and Lens, S.M.A. (2009). Sensing chromosome bi-orientation by spatial separation of aurora B kinase from kinetochore substrates. Science (New York, NY) *323*, 1350-1353.

Liu, D., Vleugel, M., Backer, C.B., Hori, T., Fukagawa, T., Cheeseman, I.M., and Lampson, M.A. (2010). Regulated targeting of protein phosphatase 1 to the outer kinetochore by KNL1 opposes Aurora B kinase. The Journal of cell biology *188*, 809-820.

Martin, D.D., Xu, M.Q., and Evans, T.C. (2001). Characterization of a naturally occurring trans-splicing intein from Synechocystis sp PCC6803. Biochemistry *40*, 1393-1402.

Martineau-Thuillier, S., Andreassen, P.R., Margolis, R.L. (1998). Colocalization of TD-60 and INCENP throughout G2 and mitosis: evidence for their possible interaction in signalling cytokinesis. Chromosoma *107*, 461--470.

Mitchison, T.J. (2012). The proliferation rate paradox in antimitotic chemotherapy. Molecular biology of the cell *23*, 1--6.

Modugno, M., Casale, E., Soncini, C., Rosettani, P., Colombo, R., Lupi, R., Rusconi, L., Fancelli, D., Carpinelli, P., Cameron, A.D., Isacchi, A., Moll, J. (2007). Crystal structure of the T315I Abl mutant in complex with the aurora kinases inhibitor PHA-739358. Cancer research *67*, 7987--7990.

Nakahara, T., Takeuchi, M., Kinoyama, I., Minematsu, T., Shirasuna, K., Matsuhisa, A., Kita, A., Tominaga, F., Yamanaka, K., Kudoh, M*., et al.* (2007). YM155, a novel small-molecule survivin suppressant, induces regression of established human hormone-refractory prostate tumor xenografts. Cancer research *67*, 8014-8021.

Niedzialkowska, E., Wang, F., Porebski, P.J., Minor, W., Higgins, J.M., Stukenberg, P.T. (2012). Molecular basis for phosphospecific recognition of histone H3 tails by Survivin paralogues at inner centromeres. Molecular biology of the cell *23*, 1457--1466.

Nishimura, Y., and Yonemura, S. (2006). Centralspindlin regulates ECT2 and RhoA accumulation at the equatorial cortex during cytokinesis. Journal of cell science *119*, 104--114.

Norden, C., Mendoza, M., Dobbelaere, J., Kotwaliwale, C.V., Biggins, S., Barral, Y. (2006). The NoCut pathway links completion of cytokinesis to spindle midzone function to prevent chromosome breakage. Cell *125*, 85--98.

Nozawa, R.-S., Nagao, K., Masuda, H.T., Iwasaki, O., Hirota, T., Nozaki, N., Kimura, H., Obusde, C. (2010). Human POGZ modulates dissociation of HP1alpha from mitotic chromosome arms through Aurora B activation. Nature cell biology *12*, 719--727.

Ozlü, N., Monigatti, F., Renard, B.Y., Field, C.M., Steen, H., Mitchison, T.J., Steen, J.J. (2010). Binding partner switching on microtubules and aurora-B in the mitosis to cytokinesis transition. Molecular \& cellular proteomics : MCP *9*, 336--350.

Pannone, G., Hindi, S.A.H., Santoro, A., Sanguedolce, F., Rubini, C., Cincione, R.I., De Maria, S., Tortorella, S., Rocchetti, R., Cagiano, S*., et al.* Aurora B expression as a prognostic indicator and possible therapeutic target in oral

squamous cell carcinoma. International journal of immunopathology and pharmacology *24*, 79-88.

Plescia, J., Salz, W., Xia, F., Pennati, M., Zaffaroni, N., Daidone, M.G., Meli, M., Dohi, T., Fortugno, P., Nefedova, Y.*, et al.* (2005). Rational design of shepherdin, a novel anticancer agent. Cancer cell *7*, 457-468.

Qian, J., Lesage, B., Beullens, M., Van Eynde, A., Bollen, M. (2011). PP1/Repo-man dephosphorylates mitotic histone H3 at T3 and regulates chromosomal aurora B targeting. Current biology : CB *21*, 766--773.

Ramadan, K., Bruderer, R., Spiga, F.M., Popp, O., Baur, T., Gotta, M., Meyer, H.H. (2007). Cdc48/p97 promotes reformation of the nucleus by extracting the kinase Aurora B from chromatin. Nature *450*, 1258--1262.

Rieder, C.L., Cole, R.W., Khodjakov, A., and Sluder, G. (1995). The checkpoint delaying anaphase in response to chromosome monoorientation is mediated by an inhibitory signal produced by unattached kinetochores. The Journal of cell biology *130*, 941-948.

Robbins, E., and Gonatas, N. (1964). The Ultrastructure of a Mammalian Cell During the Mitotic Cycle. The Journal of cell biology *21*, 429--463.

Roehrl, M.H.A., Kang, S., Aramburu, J., Wagner, G., Rao, A., and Hogan, P.G. (2004). Selective inhibition of calcineurin-NFAT signaling by blocking protein-protein interaction with small organic molecules. Proceedings of the National Academy of Sciences of the United States of America *101*, 7554-7559.

Ruchaud, S., Carmena, M., and Earnshaw, W.C. (2007). Chromosomal passengers: conducting cell division. Nature reviews Molecular cell biology *8*, 798-812.

Sampath, S.C., Ohi, R., Leismann, O., Salic, A., Pozniakovski, A., and Funabiki, H. (2004). The chromosomal passenger complex is required for chromatin-induced microtubule stabilization and spindle assembly. Cell *118*, 187-202.

Santaguida, S., and Musacchio, A. (2009). The life and miracles of kinetochores. The EMBO journal *28*, 2511--2531.

Sasai, K., Katayama, H., Stenoien, D.L., Fujii, S., Honda, R., Kimura, M., Okano, Y., Tatsuka, M., Suzuki, F., Nigg, E.A.*, et al.* (2004). Aurora-C kinase is a novel chromosomal passenger protein that can complement Aurora-B kinase function in mitotic cells. Cell motility and the cytoskeleton *59*, 249-263.

Saurin, A.T., van der Waal, M.S., Medema, R.H., Lens, S.M.A., and Kops, G.J.P.L. (2011). Aurora B potentiates Mps1 activation to ensure rapid checkpoint establishment at the onset of mitosis. Nature communications *2*, 316-316.

Schrödinger, L. (2010). The PyMOL Molecular Graphics System.

Scott, C.P., Abel-Santos, E., Jones, A.D., and Benkovic, S.J. (2001). Structural requirements for the biosynthesis of backbone cyclic peptide libraries. Chemistry & biology *8*, 801-815.

Scott, C.P., Abel-Santos, E., Wall, M., Wahnon, D.C., and Benkovic, S.J. (1999). Production of cyclic peptides and proteins in vivo. Proceedings of the National Academy of Sciences of the United States of America *96*, 13638-13643.

Sessa, F., Mapelli, M., Ciferri, C., Tarricone, C., Areces, L.B., Schneider, T.R., Stukenberg, P.T., and Musacchio, A. (2005). Mechanism of Aurora B activation by INCENP and inhibition by hesperadin. Molecular cell *18*, 379-391.

Shah, N.P., Nicoll, J.M., Nagar, B., Gorre, M.E., Paquette, R.L., Kuriyan, J., Sawyers, C.L. (2002). Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. Cancer Cell *2*, 117--125.

Sillerud, L.O., and Larson, R.S. (2005). Design and structure of peptide and peptidomimetic antagonists of protein-protein interaction. Current protein & peptide science *6*, 151-169.

Smith, S.L., Bowers, N.L., Betticher, D.C., Gautschi, O., Ratschiller, D., Hoban, P.R., Booton, R., Santibáñez-Koref, M.F., and Heighway, J. (2005). Overexpression of aurora B kinase (AURKB) in primary non-small cell lung carcinoma is frequent, generally driven from one allele, and correlates with the level of genetic instability. British journal of cancer *93*, 719-729.

Steigemann, P., Wurzenberger, C., Schmitz, M.H.A., Held, M., Guizetti, J., Maar, S., and Gerlich, D.W. (2009). Aurora B-mediated abscission checkpoint protects against tetraploidization. Cell *136*, 473-484.

Sudakin, V., Chan, G.K., Yen, T.J. (2001). Checkpoint inhibition of the APC/C in HeLa cells is mediated by a complex of BUBR1, BUB3, CDC20, and MAD2. The Journal of cell biology *154*, 925--936.

Sugiyama, K., Sugiura, K., Hara, T., Sugimoto, K., Shima, H., Honda, K., Furukawa, K., Yamashita, S., Urano. T. (2002). Aurora-B associated protein phosphatases as negative regulators of kinase activation. Oncogene *21*, 3103--3111.

Symmans, W.F., Volm, M.D., Shapiro, R.L., Perkins, A.B., Kim, A.Y., Demaria, S., Yee, H.T., McMullen, H., Oratz, R., Klein, P., Formenti, S.C., Muggia, F. (2000). Paclitaxel-induced Apoptosis and Mitotic Arrest Assessed by Serial Fine-Needle Aspiration: Implications for Early Prediction of Breast Cancer Response to Neoadjuvant Treatment. Clin Cancer Res *6*, 4610--4617.

Szewczuk, Z., Gibbs, B.F., Yue, S.Y., Purisima, E.O., Konishi, Y. (1992). Conformationally restricted thrombin inhibitors resistant to proteolytic digestion. Biochemistry *31*, 9132--9140.

Takahashi, T., Futamura, M., Yoshimi, N., Sano, J., Katada, M., Takagi, Y., Kimura, M., Yoshioka, T., Okano, Y., and Saji, S. (2000). Centrosomal kinases, HsAIRK1 and HsAIRK3, are overexpressed in primary colorectal cancers. Japanese journal of cancer research : Gann *91*, 1007-1014.

Tanaka, T.U., Rachidi, N., Janke, C., Pereira, G., Galova, M., Schiebel, E., Stark, M.J.R., and Nasmyth, K. (2002). Evidence that the Ipl1-Sli15 (Aurora kinase-INCENP) complex promotes chromosome bi-orientation by altering kinetochore-spindle pole connections. Cell *108*, 317-329.

Tavassoli, A., Lu, Q., Gam, J., Pan, H., Benkovic, S.J., and Cohen, S.N. (2008). Inhibition of HIV budding by a genetically selected cyclic peptide targeting the Gag-TSG101 interaction. ACS chemical biology *3*, 757-764.

Terada, Y., Tatsuka, M., Suzuki, F., Yasuda, Y., Fujita, S., and Otsu, M. (1998). AIM-1: a mammalian midbody-associated protein required for cytokinesis. The EMBO journal *17*, 667-676.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). Clustal-W - Improving the sensitivity of progressive multiple sequence alignment through sequence weightin, position-specific gap penalties and weight matrix choice. Nucleic Acids Research *22*, 4673-4680.

Trabi, M., and Craik, D.J. (2002). Circular proteins--no end in sight. Trends in biochemical sciences *27*, 132-138.

Udugamasooriya, D.G.a.S.M.R. (2008). Conformational constraint in protein ligand design and the inconsistency of binding entropy. Biopolymers *89*, 653--667.

Uhlmann, F., Lottspeich, F., Nasmyth, K. (1999). Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1. Nature *400*, 37--42.

Vagnarelli, P., Ribeiro, S., Sennels, L., Sanchez-Pulido, L., de Lima Alves, F., Verheyen, T., Kelly, D.A., Ponting, C.P., Rappsilber, J., Earnshaw, W.C. (2011). Repo-Man coordinates chromosomal reorganization with nuclear envelope reassembly during mitotic exit. Developmental cell *21*, 328--342.

van der Waal, M.S., Hengeveld, R.C., van der Horst, A., Lens, S.M. (2012). Cell division control by the Chromosomal Passenger Complex. Experimental cell research *318*, 1407--1420.

Verdecia, M.A., Huang, H., Dutil, E., Kaiser, D.A., Hunter, T., Noel, J.P. (2000). Structure of the human anti-apoptotic protein survivin reveals a dimeric arrangement. Nature structural biology *7*, 602--608.

Vivès, E., Brodin, P., and Lebleu, B. (1997). A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. The Journal of biological chemistry *272*, 16010-16017.

Wang, F., Dai, J., Daum, J.R., Niedzialkowska, E., Banerjee, B., Stukenberg, P.T., Gorbsky, G.J., and Higgins, J.M.G. (2010). Histone H3 Thr-3 phosphorylation by Haspin positions Aurora B at centromeres in mitosis. Science (New York, NY) *330*, 231-235.

Wang, X., Zhou, Y.X., Qiao, W., Tominaga, Y., Ouchi, M., Ouchi, T., Deng, C.X. (2006). Overexpression of aurora kinase A in mouse mammary epithelium induces genetic instability preceding mammary tumor formation. Oncogene *25*, 7148--7158.

Warbrick, E., Lane, D.P., Glover, D.M., and Cox, L.S. (1995). A small peptide inhibitor of DNA replication defines the site of interaction between the cyclin-dependent kinase inhibitor p21WAF1 and proliferating cell nuclear antigen. Current Biology *5*, 275-282.

Welburn, J.P., Vleugel, M., Liu, D., Yates, J.R. 3rd, Lampson, M.A., Fukagawa, T., Cheeseman, I.M. (2010). Aurora B phosphorylates spatially distinct targets to differentially regulate the kinetochore-microtubule interface. Molecular cell *38*, 383--392.

Wells, J.A., and McClendon, C.L. (2007). Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. Nature *450*, 1001-1009.

Wheatley, S.P., Kandels-Lewis, S.E., Adams, R.R., Ainsztein, A.M., Earnshaw, W.C. (2001). INCENP binds directly to tubulin and requires dynamic microtubules to target to the cleavage furrow. Experimental cell research *262*, 122--127.

Wild, C., Oas, T., McDanal, C., Bolognesi, D., and Matthews, T. (1992). A synthetic peptide inhibitor of human immunodeficiency virus replication: correlation between solution structure and viral inhibition. Proceedings of the National Academy of Sciences of the United States of America *89*, 10537-10541.

Wilkinson, R.W., Odedra, R., Heaton, S.P., Wedge, S.R., Keen, N.J., Crafter, C., Foster, J.R., Brady, M.C., Bigley, A., Brown, E*., et al.* (2007). AZD1152, a selective inhibitor of Aurora B kinase, inhibits human tumor xenograft growth by inducing apoptosis. Clinical cancer research : an official journal of the American Association for Cancer Research *13*, 3682-3688.

Woodman, R., Yeh, J.T.H., Laurenson, S., and Ko Ferrigno, P. (2005). Design and validation of a neutral protein scaffold for the presentation of peptide aptamers. Journal of molecular biology *352*, 1118-1133.

Wu, H., Hu, Z., and Liu, X.Q. (1998). Protein trans-splicing by a split intein encoded in a split DnaE gene of Synechocystis sp. PCC6803. Proceedings of the National Academy of Sciences of the United States of America *95*, 9226-9231.

Wu, H., Xu, M.Q., Liu, X.Q. (1998). Protein trans-splicing and functional mini-inteins of a cyanobacterial dnaB intein. Biochimica et biophysica acta *1387*, 422--432.

Xu, Z., Ogawa, H., Vagnarelli, P., Bergmann, J.H., Hudson, D.F., Ruchaud, S., Fukagawa, T., Earnshaw, W.C., and Samejima, K. (2009). INCENP-aurora B interactions modulate kinase activity and chromosome passenger complex localization. The Journal of cell biology *187*, 637-653.

Xu, Z., Vagnarelli, P., Ogawa, H., Samejima, K., and Earnshaw, W.C. (2010). Gradient of increasing Aurora B kinase activity is required for cells to execute mitosis. The Journal of biological chemistry *285*, 40163-40170.

Yamagishi, Y., Honda, T., Tanno, Y., Watanabe, Y. (2010). Two histone marks establish the inner centromere and chromosome bi-orientation. Science (New York, NY) *330*, 239--243.

Yan, X., Cao, L., Li, Q., Wu, Y., Zhang, H., Saiyin, H., Liu, X., Zhang, X., Shi, Q., and Yu, L. (2005). Aurora C is directly associated with Survivin and required for cytokinesis. Genes to cells : devoted to molecular & cellular mechanisms *10*, 617-626.

Yang, Z., Kenny, A.E., Brito, D.A., and Rieder, C.L. (2009). Cells satisfy the mitotic checkpoint in Taxol, and do so faster in concentrations that stabilize syntelic attachments. The Journal of cell biology *186*, 675-684.

Yasui, Y., Urano, T., Kawajiri, A., Nagata, K.-i., Tatsuka, M., Saya, H., Furukawa, K., Takahashi, T., Izawa, I., and Inagaki, M. (2004). Autophosphorylation of a newly identified site of Aurora-B is indispensable for cytokinesis. The Journal of biological chemistry *279*, 12997-13003.

Yuan, J., Krämer, A., Eckerdt, F., Kaufmann, M., and Strebhardt, K. (2002). Efficient internalization of the polo-box of polo-like kinase 1 fused to an Antennapedia peptide results in inhibition of cancer cell proliferation. Cancer research *62*, 4186-4190.

Yue, Z., Carvalho, A., Xu, Z., Yuan, X., Cardinale, S., Ribeiro, S., Lai, F., Ogawa, H., Gudmundsdottir, E., Gassman, R., Morrison, C.G., Ruchaud, S., Earnshaw, W.C. (2008). Deconstructing Survivin: comprehensive genetic analysis of Survivin function by conditional knockout in a vertebrate cell line. The Journal of cell biology *183*, 279--296.

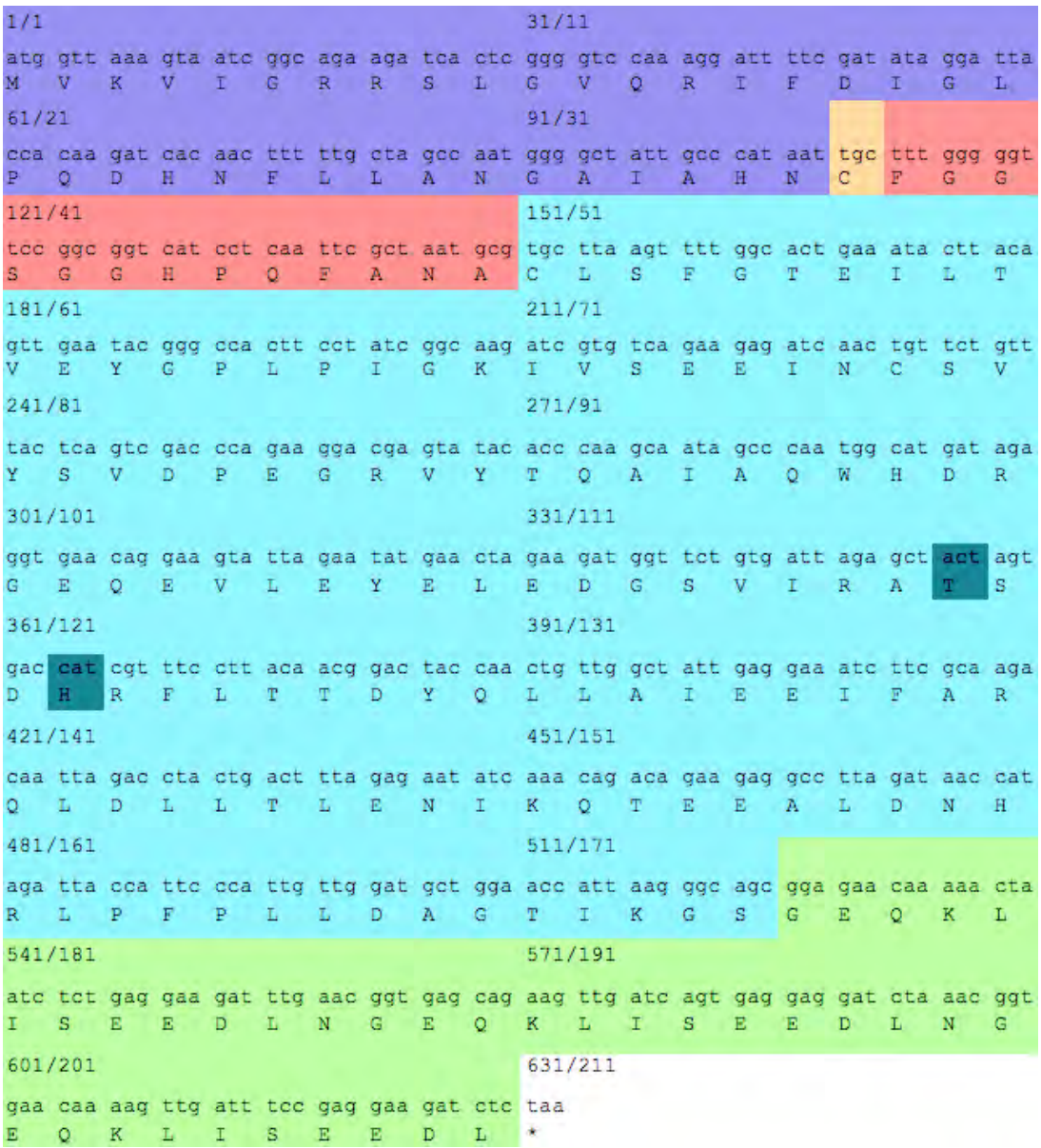
Zettler, J., Schütz, V., and Mootz, H.D. (2009). The naturally split Npu DnaE intein exhibits an extraordinarily high rate in the protein trans-splicing reaction. FEBS letters *583*, 909-914.

Zhao, Z., Ma, X., Li, L., Zhang, W., Ping, S., Xu, M.Q., Lin, M. (2010). Protein cyclization enhanced thermostability and exopeptidase-resistance of green fluorescent protein. Journal of microbiology and biotechnology *20*, 460--466.

Zutshi, R., Brickner, M., and Chmielewski, J. (1998). Inhibiting the assembly of protein-protein interfaces. Current opinion in chemical biology *2*, 62-66.

APPENDIX I – SICLOPPS TEST CONSTRUCT

Sequence of the SICLOPPS test construct used in prelimary experiments and library construction synthesized by DNA 2.0 (Menlo Park, California, USA). Colour-coding is as follows: blue = C-terminal intein; cyan = N-terminal intein; green = myc tag; orange = CP nucleophile (position 1), red = CP variable region. Darkened boxes indicate residues critical for protein splicing which, when mutated to alanine, impair the process.



APPENDIX II – INCENP INBOX LIBRARY: SEQUENCES & PRIMERS

Amino acid sequences of the human INCENP IN-box CP library clones and of the reverse primers used to generate these. The sequence indicated is that of the variable residues – the invariant cysteine present at position 1 is not shown. The first half of the table shows the sequence of all sublibraries, in ascending length, centered on residues W845 and F881. The core library spanning the INbox is covered by clones E9 through to A1.2 – this excludes the library members covering either W845 or F881. The variable region within each primer is highlighted.

The amino acid sequence of the human INCENP INbox as follows:

DEAHPRKPIPT**W**ARGTPLSQAIIHQYYHPPNLLELFGTILPLDLEDI**F**KKSKPRYHKRTSS

Universal primer		**Primer Sequence (5'-3')**
Reverse		CCGAATTCGGATCCATGGTTAAAGTAATCGGC
Forward		GCCCGTATTCAACTGTAAGTATTTCAGTGCCAAAA CTTAAGCA(*variable*)GCAATTATGGGCAATAGCC
Construct	**Peptide variable region**	**Primer Sequence variable region (5'-3')**
A1	IPTW	CCAGGTAGGGAT
A2	PTWA	GGCCCAGGTAGG
A3	TWAR	CCTGGCCCAGGT
A4	WARG	GCCCCTGGCCC
A5	PIPTW	CCAGGTAGGGATAGG
A6	IPTWA	GGCCCAGGTAGGGAT
A7	PTWAR	CCTGGCCCAGGTAGG
A8	TWARG	GCCCCTGGCCCAGGT
A9	WARGT	GGTGCCCCTGGCCCA
A10	LEDIF	GAAGATGTCCTCCAG
A11	EDIFK	CTTGAAGATGTCCTC
A12	DIFKS	GCTCTTGAAGATGTC
B1	IFKSK	CTTGCTCTTGAAGAT
B2	FKSKP	AGGCTTGCTCTTGAA
B3	KPIPTW	CCAGGTAGGGATAGGCTT
B4	PIPTWA	GGCCCAGGTAGGGATAGG
B5	IPTWAR	CCTGGCCCAGGTAGGGAT

B6	PTWARG	GCCCCTGGCCCAGGTAGG
B7	TWARGT	GGTGCCCCTGGCCCAGGT
B8	WARGTP	AGGGGTGCCCCTGGCCCA
B9	DLEDIF	GAAGATGTCCTCCAGGTC
B10	LEDIFK	CTTGAAGATGTCCTCCAG
B11	EDIFKK	CTTCTTGAAGATGTCCTC
B12	DIFKKS	GCTCTTCTTGAAGATGTC
C1	IFKKSK	CTTGCTCTTCTTGAAGAT
C2	FKKSKP	AGGCTTGCTCTTCTTGAA
C3	RKPIPTW	CCAGGTAGGGATAGGCTTCCT
C4	KPIPTWA	GGCCCAGGTAGGGATAGGCTT
C5	PIPTWAR	CCTGGCCCAGGTAGGGATAGG
C6	IPTWARG	GCCCCTGGCCCAGGTAGGGAT
C7	PTWARGT	GGTGCCCCTGGCCCAGGTAGG
C8	TWARGTP	AGGGGTGCCCCTGGCCCAGGT
C9	WARGTPL	CAGAGGGGTGCCCCTGGCCCA
C10	LDLEDIF	GAAGATGTCCTCCAGGTCCAG
C11	DLEDIFK	CTTGAAGATGTCCTCCAGGTC
C12	LEDIFKK	CTTCTTGAAGATGTCCTCCAG
D1	EDIFKKS	GCTCTTCTTGAAGATGTCCTC
D2	DIFKKSK	CTTGCTCTTCTTGAAGATGTC
D3	IFKKSKP	AGGCTTGCTCTTCTTGAAGAT
D4	FKKSKPR	CCTAGGCTTGCTCTTCTTGAA
D5	PRKPIPTW	CCAGGTAGGGATAGGCTTCCTAGGG
D6	RKPIPTWA	GGCCCAGGTAGGGATAGGCTTCCT
D7	KPIPTWAR	CCTGGCCCAGGTAGGGATAGGCTT
D8	PIPTWARG	GCCCCTGGCCCAGGTAGGGATAGG
D9	IPTWARGT	GGTGCCCCTGGCCCAGGTAGGGAT
D10	PTWARGTP	AGGGGTGCCCCTGGCCCAGGTAGG
D11	TWARGTPL	CAGAGGGGTGCCCCTGGCCCAGGT
D12	WARGTPLS	GCTCAGAGGGGTGCCCCTGGCCCA
E1	PLDLEDIF	GAAGATGTCCTCCAGGTCCAGAGG
E2	LDLEDIFK	CTTGAAGATGTCCTCCAGGTCCAG
E3	DLEDIFKK	CTTCTTGAAGATGTCCTCCAGGTC
E4	LEDIFKKS	GCTCTTCTTGAAGATGTCCTCCAG
E5	EDIFKKSK	CTTGCTCTTCTTGAAGATGTCCTC
E6	DIFKKSKP	AGGCTTGCTCTTCTTGAAGATGTC
E7	IFKKSKPR	CCTAGGCTTGCTCTTCTTGAAGAT
E8	FKKSKPRY	GTACCTAGGCTTGCTCTTCTTGAA
E9	DEAHPRK	CTTCCTAGGGTGGGCCTCGTC
E10	EAHPRKP	AGGCTTCCTAGGGTGGGCCTC
E11	AHPRKPI	GATAGGCTTCCTAGGGTGGGC
E12	HPRKPIP	AGGGATAGGCTTCCTAGGGTG
F1	PRKPIPT	GGTAGGGATAGGCTTCCTAGG

F2	ARGTPLS	GCTCAGAGGGGTGCCCCTGGC
F3	RGTPLSQ	CTGGCTCAGAGGGGTGCCCCT
F4	GTPLSQA	GGCCTGGCTCAGAGGGGTGCC
F5	TPLSQAI	GATGGCCTGGCTCAGAGGGGT
F6	PLSQAII	GATGATGGCCTGGCTCAGAGG
F7	LSQAIIH	GTGGATGATGGCCTGGCTCAG
F8	SQAIIHQ	CTGGTGGATGATGGCCTGGCT
F9	QAIIHQY	GTACTGGTGGATGATGGCCTG
F10	AIIHQYY	GTAGTACTGGTGGATGATGGC
F11	IIHQYYH	GTGGTAGTACTGGTGGATGAT
F12	IHQYYHP	AGGGTGGTAGTACTGGTGGAT
G1	HQYYHPP	AGGAGGGTGGTAGTACTGGTG
G2	QYYHPPN	GTTAGGAGGGTGGTAGTACTG
G3	YYHPPNL	CAGGTTAGGAGGGTGGTAGTA
G4	YHPPNLL	CAGCAGGTTAGGAGGGTGGTA
G5	HPPNLLE	CTCCAGCAGGTTAGGAGGGTG
G6	PPNLLEL	CAGCTCCAGCAGGTTAGGAGG
G7	PNLLELF	GAACAGCTCCAGCAGGTTAGG
G8	NLLELFG	GCCGAACAGCTCCAGCAGGTT
G9	LLELFGT	GGTGCCGAACAGCTCCAGCAG
G10	LELFGTI	GATGGTGCCGAACAGCTCCAG
G11	ELFGTIL	CAGGATGGTGCCGAACAGCTC
G12	LFGTILP	AGGCAGGATGGTGCCGAACAG
H1	FGTILPL	CAGAGGCAGGATGGTGCCGAA
H2	GTILPLD	GTCCAGAGGCAGGATGGTGCC
H3	TILPLDL	CAGGTCCAGAGGCAGGATGGT
H4	ILPLDLE	CTCCAGGTCCAGAGGCAGGAT
H5	LPLDLED	GTCCTCCAGGTCCAGAGGCAG
H6	PLDLEDI	GATGTCCTCCAGGTCCAGAGG
H7	KKSKPRY	GTACCTAGGCTTGCTCTTCTT
H8	KSKPRYH	GTGGTACCTAGGCTTGCTCTT
H9	SKPRYHK	CTTGTGGTACCTAGGCTTGCT
H10	KPRYHKR	CCTCTTGTGGTACCTAGGCTT
H11	PRYHKRT	GGTCCTCTTGTGGTACCTAGG
H12	RYHKRTS	GCTGGTCCTCTTGTGGTACCT
A1.2	YHKRTSS	GCTGCTGGTCCTCTTGTGGTA